TESIS – SS14 2501

PEMODELAN REGRESI DERET FOURIER DAN SPLINE *TRUNCATED* DALAM REGRESI NONPARAMETRIK MULTIVARIABEL (APLIKASI: DATA KEMISKINAN DI PROVINSI PAPUA)

NI PUTU AYU MIRAH MARIATI
NRP.1313 201 019

Dosen Pembimbing:
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.

PROGRAM MAGISTER
JURUSAN STATISTIKA
PROGRAM PASCA SARJANA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

THESIS – SS14 2501

MODELLING FOURIER SERIES AND SPLINE *TRUNCATED* IN NONPARAMETRIC REGRESSION MULTIVARIABLE (APPLICATIONS: DATA OF POVERTY IN PAPUA)

NI PUTU AYU MIRAH MARIATI
NRP.1313 201 019

SUPERVISOR:
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.

PROGRAM OF MAGISTER
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

PEMODELAN REGRESI DERET FOURIER DAN SPLINE *TRUNCATED* DALAM REGRESI NONPARAMETRIK MULTIVARIABEL (APLIKASI: DATA KEMISKINAN DI PROVINSI PAPUA)

Tesis ini disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si.)
di
Institut Teknologi Sepuluh Nopember Surabaya

Oleh:

Ni Putu Ayu Mirah Mariati
NRP. 1313 201 019

Tanggal Ujian : 16 Januari 2015
Periode Wisuda: Maret 2015

Disetujui oleh:

1. **Prof. Dr.Drs. I Nyoman Budiantara, M.Si.** **(Pembimbing)**
   NIP. 19650603 198903 1 003

Ratnasari -

2. **Dr. Vita Ratnasari, M.Si.** **(Penguji)**
   NIP. 19700910 199702 2 001

3. **Dr. Santi Wulan Purnami, M.Si.** **(Penguji)**
   NIP. 19720923 199803 2 001

Direktur Program Pascasarjana

KEMENTERIAN INSTITUT TEKNOLOGI SEPULUH NOPEMBER DAN KEBUDAYAAN PROGRAM PASCASARJANA

Prof.Dr.Ir.Adi Soeprijanto, M.T.
NIP. 19640405 199002 1 001

KATA PENGANTAR

Om Swastiastu,

Puja dan puji syukur penulis panjatkan kehadirat Ida Sang Hyang Widhi Wasa/Tuhan Yang Maha Esa atas berkat dan kelancaran yang telah diberikan, sehingga penulis dapat menyelesaikan Tesis yang berjudul: **"Pemodelan Regresi Deret Fourier dan Spline *Truncated* dalam Regresi Nonparametrik Multivariabel (Aplikasi: Data Kemiskinan di Provinsi Papua)"**. Tesis ini disusun sebagai salah satu syarat untuk menyelesaikan Studi Magister Statistika pada Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember Surabaya.

Seiring dengan puji dan rasa syukur, penulis menyampaikan ucapan terima kasih kepada berbagai pihak yang telah berjasa selama penulis menempuh pendidikan program Magister (S2) di Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember Surabaya. Banyak pihak yang telah dan selalu memberikan saran, bantuan, kritik, semangat, dan motivasi dalam pembuatan laporan tesis ini. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Bapak (I Wayan Sudiarsa) dan Mama (Ni Ketut Martini) yang selalu menyemangati dan memberi motivasi penulis selama pembuatan tesis ini. Kalian orangtua terbaik dan tersayang bagi penulis. Terimakasih untuk semuanya.
2. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si., selaku dosen pembimbing, yang selalu sabar memberikan bimbingan dan memotivasi saya selama proses penulisan tesis ini. Beliau sudah seperti sosok Bapak bagi saya di kampus ITS. Terima kasih atas segala bimbingan.
3. Ibu Dr. Vita Ratnasari, S.Si, M.Si. dan Ibu Dr. Santi Wulan Purnami, M.Si. selaku dosen penguji yang telah memberikan masukan dan saran dalam penyempurnaan tesis ini. Terima kasih atas segala bimbingannya.
4. Bapak Dr. Muhammad Mashuri, M.T selaku ketua Jurusan Statistika FMIPA ITS.

5. Semua dosen yang telah banyak memberikan bantuan dari awal kuliah sampai penulisan tesis.
6. Made Ary Januardana, untuk segalanya. Kasih sayang tulus dan kesetiaan yang telah diberikan. Terutama untuk selalu memberikan kepercayaan serta pencerahan hingga penulis bisa menyelesaikan tesis.
7. Adik-adikku tersayang (Ni Made Sukma Sanjiwani dan Ni Komang Indra Mahayani) serta seluruh keluarga besar penulis yang selalu memberikan semangat. Terimakasih untuk semuanya.
8. BPPDN Dikti yang telah memberikan beasiswa kepada peneliti selama mengikuti Program Magister, pada Program Pasca Sarjana, Jurusan Statistika, Institut Teknologi Sepuluh Nopember (ITS) Surabaya.
9. Teman-teman seperjuangan dari daerah asal (Bali) Kak Amik, Safitri, Kak Nanik dan Kak pal. Terima kasih atas semua semangat dan dukungan.
10. Teman sebimbingan Fariz, Mbak Lilis dan Kak Bobby begitu juga untuk Vellin dan Sulis.
11. Teman – teman seperjuangan penulis (S2 angkatan 2013) terimakasih atas segala dukungan dalam pembuatan tesis ini.
12. Semua pihak yang tidak bisa disebutkan satu per satu oleh penulis, terima kasih atas semua dukungan dan semangat moral yang diberikan.

Penulis menyadari bahwa penyusunan tesis ini masih jauh dari kesempurnaan. Karena sebagai manusia biasa tidak luput dari berbagai kekurangan atau keterbatasannya. Tesis ini juga tak luput dari pepatah, “Tak ada gading yang tak retak”. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang konstrutif dan membangun agar tesis ini menjadi lebih baik.

Om Shanti, Shanti, Shanti Om

Surabaya, Januari 2015

Ni Putu Ayu Mirah Mariati

PEMODELAN REGRESI DERET FOURIER DAN SPLINE TRUNCATED DALAM REGRESI NONPARAMETRIK MULTIVARIABEL (APLIKASI: DATA KEMISKINAN DI PROVINSI PAPUA)

Nama : **Ni Putu Ayu Mirah Mariati**
NRP : **1313 201 019**
Jurusan : **Statistika FMIPA-ITS**
Dosen Pembimbing : **Prof. Dr. Drs. I Nyoman Budiantara, M.Si**

ABSTRAK

Analisa regresi digunakan untuk menyelidiki pola hubungan variabel dependen dengan variabel independen. Hal ini dapat dilakukan dengan dua pendekatan. Pendekatan yang paling umum dan sering digunakan adalah pendekatan parametrik. Pendekatan parametrik mengasumsikan bentuk model mengikuti pola tertentu. Apabila tidak ada informasi apapun tentang bentuk dari fungsi regresi, maka pendekatan yang digunakan adalah pendekatan regresi nonparametrik. Terdapat beberapa pendekatan dalam estimasi kurva regresi nonparametrik diantaranya spline truncated dan deret fourier. Kelebihan spline truncated adalah dapat menggambarkan perubahan pola perilaku dari fungsi pada sub interval tertentu dan deret fourier baik digunakan untuk menjelaskan kurva yang polanya berulang. Tujuan penelitian ini adalah mengkaji estimasi regresi nonparametrik Spline *Truncated* begitu pula estimasi regresi Deret Fourier dan menerapkan pada data kemiskinan di provinsi Papua.

Estimasi kurva regresi nonparametrik Spline *Truncated* adalah:

$$\hat{\mu}\left(x_{1i}, x_{2i}, ..., x_{pi}\right) = \sum_{j=1}^{p}\sum_{v=1}^{m}\hat{\beta}_{vj}x_{ji}^{v} + \sum_{j=1}^{p}\sum_{k=1}^{r}\hat{\beta}_{j(k+m)}\left(x_{ji} - K_{jk}\right)_{+}^{m}$$

Estimator parameter regresi Spline *Truncated* diberikan oleh $\underset{\sim}{\hat{\beta}} = \left(\underset{\sim}{\hat{\beta}}_1', ..., \underset{\sim}{\hat{\beta}}_p'\right)'$,

dengan $\underset{\sim}{\hat{\beta}}_1 = \left(\beta_{11}, ..., \beta_{m1}, \beta_{1(1+m)}, ..., \beta_{1(r+m)}\right)', ..., \underset{\sim}{\hat{\beta}}_p = \left(\beta_{1p}, ..., \beta_{mp}, \beta_{p(1+m)}, ..., \beta_{p(r+m)}\right)'$.

Estimasi kurva regresi nonparametrik Deret Fourier adalah:

$$\hat{\mu}\left(x_{1i}, x_{2i}, ..., x_{qi}\right) = \sum_{j=1}^{q}\left(\hat{b}_j x_{ji} + \frac{1}{2}\hat{\alpha}_{0j} + \sum_{k=1}^{K}\hat{\alpha}_{kj}\cos kx_{ji}\right)$$

Estimator parameter regresi Deret Fourier diberikan oleh

$$\hat{\beta} = \left[\hat{b}_1 \quad \frac{1}{2}\hat{\alpha}_{01} \quad \hat{\alpha}_{11} \quad \cdots \quad \hat{\alpha}_{K1} \quad \vdots \quad \cdots \quad \vdots \quad \hat{b}_q \quad \frac{1}{2}\hat{\alpha}_{0q} \quad \hat{\alpha}_{1q} \quad \cdots \quad \hat{\alpha}_{Kq}\right]'.$$

Provinsi Papua menduduki peringkat pertama presentase penduduk miskin di Indonesia. Model regresi nonparametrik Spline *Truncated* terbaik untuk model

kemiskinan di Papua adalah dengan tiga titik knot. Model Spline *Truncated* terbaik diberikan oleh:

$$\begin{aligned}\hat{y} \ = \ & 145{,}11+0{,}64\,x_1 - 0{,}83\,(x_1 - 61{,}76)_+ - 0{,}39\,(x_1 - 73{,}48)_+ + \ 218{,}58\,(x_1 - 98{,}38)_+ + \\ & -8{,}90\,x_2 + 13{,}97\,(x_2 - 6{,}41)_+ - 27{,}72\,(x_2 - 7{,}84)_+ - 2182{,}53\,(x_2 - 10{,}88)_+ + \\ & 0{,}11\,x_3 + 0{,}41\,(x_3 - 50{,}63)_+ - 1{,}87\,(x_3 - 64{,}79)_+ + 8{,}31\,(x_3 - 94{,}89)_+ + \\ & -5{,}82\,x_4 + 9{,}44(x_4 - 53{,}73)_+ - 1{,}63(x_4 - 67{,}96)_+ - 45{,}01(x_4 - 98{,}22)_+ + \\ & 3{,}08\,x_5 - 10\,(x_5 - 65{,}79)_+ + 5{,}95\,(x_5 - 76{,}31)_+ + 70{,}24\,(x_5 \ - \ 98{,}68)_+ .\end{aligned}$$

Sedangkan untuk model regresi nonparametrik Deret Fourier terbaik adalah dengan K=3. Berikut adalah model yang terbaik berdasarkan data kemiskinan di provinsi Papua.

$$\begin{aligned}\hat{y}_i = & 16{,}88 - 0{,}27x_{1i} - 0{,}02\cos x_{1i} + 0{,}80\cos 2x_{1i} - 3{,}66\cos 3x_{1i} \\ & 3{,}69x_{2i} + -1{,}36\cos x_{2i} - 4{,}95\cos 2x_{2i} - 0{,}04\cos 3x_{2i} \\ & 0{,}41x_{3i} + 6{,}35\cos x_{3i} - 2{,}99\cos 2x_{3i} + 0{,}49\cos 3x_{3i} \\ & -0{,}32x_{4i} - 13{,}58\cos x_{4i} - 9{,}10\cos 2x_{4i} - 9{,}48\cos 3x_{4i} \\ & 0{,}14x_{5i} + 4{,}79\cos x_{5i} + 10{,}28\cos 2x_{5i} + 4{,}55\cos 3x_{5i}\end{aligned}$$

Berdasarkan pemodelan yang telah dilakukan menggunakan Spline *Truncated* dan Deret Fourier pada kemiskinan di Papua maka dapat disimpulkan bahwa metode Spline *Truncated* lebih baik. Hal ini dikarenakan nilai GCV yang minimum dan R^2 lebih besar dibandingkan dengan Deret Fourier. Dengan nilai GCV yang diperoleh yaitu 16,70 dan R^2 sebesar 98,46%.

Kata kunci: Regresi Nonparametrik, Deret Fourier, Spline *Truncated*, Data Kemiskinan

MODELLING FOURIER SERIES AND SPLINE TRUNCATED REGRESSION IN NONPARAMETRIC REGRESSION MULTIVARIABLE (APPLICATIONS: DATA OF POVERTY IN PAPUA)

Name of Student : Ni Putu Ayu Mirah Mariati
NRP : 1313 201 019
Departement : Statistics FMIPA-ITS
Supervisor : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

ABSTRACT

Regression analysis is used to investigate the relationship pattern dependent variable and independent variables. This can be done by two approaches. The most common approach and is often used is a parametric approach. Parametric approach assumes the form of a model to follow a certain pattern. If there is no any information about the shape of the regression function, the approach used is a nonparametric approach. There are several approaches to estimate such nonparametric regression spline curve and the truncated Fourier series. Excess truncated spline is able to describe the change in the pattern of behavior of the function at a particular sub-intervals and better Fourier series is used to describe a recurring pattern curve. The purpose of this study is to examine the Spline Truncated nonparametric regression estimation as well as the estimated Fourier Series regression and apply the data on poverty in Papua province.
Estimation of Spline Truncated nonparametric regression curve is:

$$\hat{\mu}\left(x_{1i}, x_{2i}, ..., x_{pi}\right) = \sum_{j=1}^{p}\sum_{v=1}^{m}\hat{\beta}_{vj}x_{ji}^{v} + \sum_{j=1}^{p}\sum_{k=1}^{r}\hat{\beta}_{j(k+m)}\left(x_{ji} - K_{jk}\right)_{+}^{m}$$

parameter estimate Spline Truncated is given by $\hat{\underset{\sim}{\beta}} = \left(\hat{\underset{\sim}{\beta}}_1', ..., \hat{\underset{\sim}{\beta}}_p'\right)'$,

with $\hat{\underset{\sim}{\beta}}_1 = \left(\beta_{11}, ..., \beta_{m1}, \beta_{1(1+m)}, ..., \beta_{1(r+m)}\right)', ..., \hat{\underset{\sim}{\beta}}_p = \left(\beta_{1p}, ..., \beta_{mp}, \beta_{p(1+m)}, ..., \beta_{p(r+m)}\right)'$.

Estimation of Fourier Series Nonparametric regression curve are:

$$\hat{\mu}\left(x_{1i}, x_{2i}, ..., x_{qi}\right) = \sum_{j=1}^{q}\left(\hat{b}_j x_{ji} + \frac{1}{2}\hat{\alpha}_{0j} + \sum_{k=1}^{K}\hat{\alpha}_{kj}\cos kx_{ji}\right)$$

parameter estimate Fourier Series regression is given by

$$\hat{\beta} = \left[\hat{b}_1 \quad \frac{1}{2}\hat{\alpha}_{01} \quad \hat{\alpha}_{11} \quad \cdots \quad \hat{\alpha}_{K1} \quad \vdots \quad \cdots \quad \vdots \quad \hat{b}_q \quad \frac{1}{2}\hat{\alpha}_{0q} \quad \hat{\alpha}_{1q} \quad \cdots \quad \hat{\alpha}_{Kq}\right]'.$$

Papua Province ranked first percentage of poor people in Indonesia. The best model of Spline Truncated nonparametric regression for models of poverty in Papua is the three point knots. Here are the best models that have been obtained are

$$\hat{y} = 145{,}11 + 0{,}64\, x_1 - 0{,}83\,(x_1 - 61{,}76)_+ - 0{,}39\,(x_1 - 73{,}48)_+ + 218{,}58\,(x_1 - 98{,}38)_+ +$$
$$-8{,}90\, x_2 + 13{,}97\,(x_2 - 6{,}41)_+ - 27{,}72\,(x_2 - 7{,}84)_+ - 2182{,}53\,(x_2 - 10{,}88)_+ +$$
$$0{,}11\, x_3 + 0{,}41\,(x_3 - 50{,}63)_+ - 1{,}87\,(x_3 - 64{,}79)_+ + 8{,}31\,(x_3 - 94{,}89)_+ +$$
$$-5{,}82\, x_4 + 9{,}44(x_4 - 53{,}73)_+ - 1{,}63(x_4 - 67{,}96)_+ - 45{,}01(x_4 - 98{,}22)_+ +$$
$$3{,}08\, x_5 - 10\,(x_5 - 65{,}79)_+ + 5{,}95\,(x_5 - 76{,}31)_+ + 70{,}24\,(x_5 - 98{,}68)_+.$$

As for nonparametric regression model is best Fourier series with K = 3. Here are the best model based on the data on poverty in the province.

$$\hat{y}_i = 16{,}88 - 0{,}27x_{1i} - 0{,}02\cos x_{1i} + 0{,}80\cos 2x_{1i} - 3{,}66\cos 3x_{1i}$$
$$3{,}69x_{2i} + -1{,}36\cos x_{2i} - 4{,}95\cos 2x_{2i} - 0{,}04\cos 3x_{2i}$$
$$0{,}41x_{3i} + 6{,}35\cos x_{3i} - 2{,}99\cos 2x_{3i} + 0{,}49\cos 3x_{3i}$$
$$-0{,}32x_{4i} - 13{,}58\cos x_{4i} - 9{,}10\cos 2x_{4i} - 9{,}48\cos 3x_{4i}$$
$$0{,}14x_{5i} + 4{,}79\cos x_{5i} + 10{,}28\cos 2x_{5i} + 4{,}55\cos 3x_{5i}$$

Based on the modeling done by using Spline Truncated and Fourier Series on poverty data in Papua, it can be concluded that the method spline Truncated is better than the Fourier series. This is because the minimum GCV and R^2 is greater than the Fourier series.With GCV value obtained is 16.70 and R^2 of 98.46%.

Keywords: Nonparametric Regression, Fourier Series, Spline Truncated, Poverty Data.

DAFTAR ISI

Halaman

(...Halaman Ini Sengaja Dikosongkan...)

DAFTAR TABEL

Halaman

(...Halaman Ini Sengaja Dikosongkan...)

DAFTAR GAMBAR

Halaman

(...Halaman Ini Sengaja Dikosongkan...)

BAB 1
PENDAHULUAN

1.1 Latar Belakang

Memodelkan satu atau lebih variabel, hal yang pertama semestinya dilakukan adalah apakah variabel tersebut secara rasional berkorelasi atau tidak. Apabila secara rasional terjadi korelasi, maka dapat dilakukan pemodelan Statistika dengan menggunakan analisis regresi. Analisis regresi merupakan salah satu analisis dalam Statistika yang digunakan untuk menyelidiki pola hubungan fungsional antara satu atau lebih variabel. Disamping itu, tujuan lain analisis regresi adalah untuk memprediksi (Budiantara, 2009). Analisis regresi dalam mengestimasi kurva regresi terdapat tiga pendekatan, yaitu pendekatan regresi parametrik, regresi nonparametrik dan regresi semiparametrik. Dalam pendekatan regresi parametrik terdapat asumsi yang sangat kuat dan kaku yaitu bentuk kurva regresi diketahui misalnya linier, kuadratik, kubik, polinomial derajat p, eksponen dan lain-lain. Disamping itu, diperlukan pengetahuan masa lalu tentang karakteristik data agar memperoleh pemodelan yang baik. Tujuan utama dalam analisis regresi adalah mencari bentuk estimasi kurva regresi. Dalam model regresi parametrik estimasi kurva regresi ekuivalen dengan estimasi terhadap parameter-parameter dalam model (Budiantara, 2009).

Berbeda dengan pendekatan regresi parametrik, dalam regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui. Kurva regresi nonparametrik hanya diasumsikan *smooth* (mulus) dalam arti termuat di dalam suatu ruang fungsi tertentu. Data diharapkan mencari sendiri bentuk estimasinya, tanpa dipengaruhi oleh faktor subyektifitas dari perancang penelitian. Dengan demikian, pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank,1988). Beberapa model regresi nonparametrik yang banyak digunakan diantaranya Kernel (Hardle,1990), Spline (Wahba, 1990; Budiantara,2009), *K-Nearest Neighbor* (Hardle,1990), Estimator Deret Fourier (Eubank, 1999), Histogram (Green dan Silverman, 1995), MARS, Deret Ortogonal, Wavelets, Neural Network. Pendekatan regresi nonparametrik yang cukup populer adalah spline. Spline merupakan potongan-potongan polinomial yang memiliki sifat tersegmen dan kontinu. Salah satu kelebihan spline adalah model ini cenderung mencari sendiri estimasi data kemanapun pola data tersebut bergerak. Kelebihan ini terjadi karena dalam spline terdapat titik-titik knot, yaitu titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data (Eubank, 1999; Budiantara,2009). Dengan titik knot ini, spline dapat memberikan fleksibilitas yang lebih baik dari pada polinomial, sehingga memungkinkan untuk menyesuaikan diri secara efektif terhadap karakteristik lokal dari suatu fungsi atau data.

Spline sebagai pendekatan pola data dikenalkan oleh Whittaker pada tahun 1923. Spline dipopulerkan oleh Schoenberg pada tahun 1942. Sedangkan spline yang didasarkan pada suatu persoalan optimasi dikembangkan oleh Reinsch pada tahun 1967 (Wahba, 1990). Pendekatan spline mempunyai suatu basis fungsi. Basis fungsi yang biasa digunakan antara lain spline *truncated* dan B-spline. Spline *Truncated* merupakan fungsi dimana terdapat perubahan pola perilaku kurva yang berbeda pada interval-interval yang berlainan. Spline adalah salah satu bentuk estimator yang juga seringkali digunakan dalam regresi nonparametrik karena mempunyai intepretasi visual yang baik, fleksibel, serta mampu menangani karakter fungsi yang bersifat mulus (Eubank, 1988; Budiantara, 2007). Kelebihan Spline adalah dapat menggambakan perubahan pola prilaku dari fungsi pada sub interval tertentu. Di samping itu juga mempunyai keunggulan dalam mengatasi pola data yang menunjukkan naik/turun yang tajam dengan bantuan titik-titik knot serta kurva yang dihasilkan relatif mulus (Eubank, 1999). Beberapa peneliti yang telah mengkaji tentang fungsi spline dalam regresi nonparametrik diantaranya Craven dan Wahba (1979) melakukan kajian tentang metode cross validation (CV) untuk memilih parameter penghalus optimal dalam estimator spline. Selain itu, Wahba (1990) mengkaji tentang metode untuk memilih parameter penghalus optimal dalam estimator spline, yaitu metode *Generalized Cross Validation* (GCV). Dalam perkembangannya, Budiantara (2000) telah melakukan kajian untuk menggeneralisasi metode GCV dari Wahba (1990) untuk estimator spline terbobot dan memperlihatkan sifat optimal asimtotik metode ini masih berlaku untuk estimator spline terbobot. Mubarak (2012) melakukan penelitian menggunakan regresi spline multivariabel untuk pemodelan kematian penderita Demam Berdarah Dengue (DBD) di Jawa Timur. Bintariningrum, (2014) meneliti tetang angka kelahiran kasar di Surabaya menggunakan regresi nonparametrik spline.

Selain itu, metode estimasi yang banyak mendapat perhatian dari beberapa peneliti regresi nonparametrik salah satunya adalah estimator Deret Fourier. Deret Fourier mulai populer pada tahun 1992 yang dipopulerkan oleh Bilodeau. Deret Fourier merupakan polinomial trigonometri yang mempunyai fleksibilitas, sehingga dapat menyesuaikan diri secara efektif terhadap sifat lokal data (Asrini dan Budiantara, 2014; Pane et.al., 2014). Deret Fourier baik digunakan untuk menjelaskan kurva yang polanya berulang (Asrini, 2012). Estimator Deret Fourier ini umumnya digunakan apabila data yang diselidiki polanya tidak diketahui dan ada kecenderungan pola berulang (terdapat pengulangan nilai respon pada nilai variabel prediktor yang berbeda) (Tripena dan Budiantara, 2007; Bilodeau, 1992). Selanjutnya Tripena dan Budiantara (2007) melakukan analisis terhadap estimator Deret Fourier yang diberikan oleh Bilodeau (1992). Tjahjono (2009) melakukan penelitian mengenai estimator Deret Fourier terbobot pada regresi nonparametrik. Selanjutnya, penelitian dilakukan oleh Semiati (2010) dan Tripena (2013) tentang estimasi Deret Fourier pada regresi nonparametrik birespon. Prahutama (2013) melakukan penelitian mengenai model regresi nonparametrik dengan pendekatan Deret Fourier pada kasus tingkat pengangguran terbuka di Jawa Timur. Sehingga berdasarkan kelebihan dari Spline *Truncated* dan Deret Fourier maka akan dilakukan perbandingan antara Spline *Truncated* dengan Deret Fourier.

Dalam banyak kasus, hubungan antar variabel tidak selalu berpola parametrik atau berpola nonparametrik saja. Menurut Eubank (1988), dalam beberapa kasus yang lain, sering terdapat hubungan antar variabel yang berpola semiparametrik. Dalam keadaan dimana bentuk kurva regresi terdiri dari komponen parametrik yang diketahui bentuk polanya dan komponen nonparametrik yang *smooth* (halus, mulus, licin). Dalam penelitian ini dilakukan metode estimasi terhadap regresi Spline *Truncated* dan Deret Fourier dengan metode *Least Square*, alasan menggunakan metode *Least Square* karena metode tersebut sederhana dan bebas dari distribusi.

Menurut Rencher (2002), regresi dapat dibedakan dari jumlah variabelnya, baik variabel respon maupun variabel prediktor yaitu regresi linier sederhana, regresi linier berganda (*multiple regression*), regresi multirespon dan regresi multivariabel. Regresi linier sederhana yaitu regresi yang terdiri dari variabel respon dan satu variabel prediktor yang polanya linier. Regresi linier berganda (*multiple regression*) jika terdiri dari satu variabel respon dengan lebih dari satu variabel prediktor dan berpola linier. Regresi dengan multivariabel tersebut jika variabel prediktor lebih dari satu. Banyak penelitian yang melakukan kajian mengenai mutivariabel salah satunya berkaitan dengan masalah sosial. Masalah yang sering dijumpai jumlah penduduk yang banyak, yang sebagian besar memiliki tingkat pendidikan yang rendah akan memicu terjadinya kemiskinan. Salah satu negara yang memiliki masalah dengan persentase kemiskinan adalah Indonesia. Indonesia merupakan negara berkembang dengan jumlah penduduk yang sangat besar, dengan jumlah penduduk terbanyak keempat didunia dan sebagian besar penduduknya memiliki tingkat pendidikan rendah. Tujuan dari pembangunan milenium yang telah disepakati oleh anggota Perserikatan Bangsa–Bangsa (PBB) dan Konferensi Tingkat Tinggi adalah Millenium Development Goals (MDGs). MDGs yang disepakati sejak tahun 1990 hingga 2015 memiliki tujuan untuk mempercepat pembangunan dan pengentasan kemiskinan. Fokus yang tersirat dari deklarasi ini adalah meningkatkan kesejahteraan manusia dalam berbagai aspek. Salah satu aspek dari kesejahteraan manusia adalah kemiskinan.

Penelitian mengenai pemodelan penduduk miskin di Jawa Timur juga pernah dilakukan oleh Damayanti (2013) dan wilayah Jawa Tengah (Prasetyawan (2010) dengan metode *Geographically Weighted Regression* (GWR). Surya (2013) juga meneliti mengenai persentase penduduk miskin di Jawa Timur menggunakan regresi Semiparametrik Spline. Selanjutnya, Ekasari (2012) melakukan pemodelan penduduk miskin di Jawa Tengah dengan menggunakan metode SEM dan GSCA. Pada tahun yang sama, Samsodin (2012) melakukan pemodelan indikator kemiskinan di Jawa Timur dengan menggunakan regresi

Spline Polynomial *Truncated* Multirespon. Penelitian mengenai jumlah penduduk miskin di Jawa Timur pernah dilakukan oleh Fadillah (2010). Metode yang digunakan adalah metode analisis regresi linier berganda. Selain itu, Pintowati (2012) meneliti mengenai Pemodelan Kemiskinan di Provinsi Jawa Timur dengan Pendekatan *Multivariate Adaptive*. Sehingga, berdasarkan hal tersebut belum ada yang meneliti penduduk miskin di Provinsi Papua.

Berdasarkan Kementrian Daerah tertinggal pada tahun 2014 Provinsi Papua merupakan provinsi urutan pertama yang merupakan daerah tertinggal di Indonesia. Dalam Rusdarti (2013) terdapat tabel peringkat presentase penduduk miskin berdasarkan provinsi di Indonesia. Pada tabel tersebut, Provinsi Papua menduduki peringkat pertama. Selain itu, Eddy (2010) melakukan penelitian mengenai pemetaan kemiskinan secara makro. Dalam penelitian tersebut didasarkan pada nilai indeks kedalaman kemiskinan karena indeks kedalaman merupakan ukuran rata-rata kesenjangan pengeluaran setiap penduduk miskin, pada penelitian tersebut terdapat peringkat Indeks Kedalaman Kemiskinan di Indonesia, Provinsi Papua merupakan urutan kedua indeks kedalaman tertinggi. Papua mempunyai indeks kedalaman diatas rata-rata indeks kedalaman di Indonesia. Berdasarkan hal tersebut pada penelitian ini akan dimodelkan presentase kemiskinan dengan menggunakan regresi Spline *Truncated* dan Deret Fourier pada salah satu provinsi di Indonesia yaitu Provinsi Papua. Selanjutnya, membandingkan antara regresi Spline *Truncated* dan Deret Fourier sehingga akan diketahui berdasarkan data kemiskinan di Provinsi Papua akan diketahui mana metode yang baik untuk digunakan.

1.2 Rumusan Masalah

Permasalahan dalam penelitian ini adalah

1. Bagaimana estimasi dari kurva regresi nonparametrik Spline *Truncated* multivariabel?
2. Bagaimana estimasi dari kurva regresi Deret Fourier dalam regresi nonparametrik multivariabel?

3. Bagaimana memodelkan penduduk miskin di Provinsi Papua dengan menggunakan regresi nonparametrik Spline *Truncated* dan Deret Fourier?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah

1. Mendapatkan estimasi dari kurva regresi nonparametrik Spline *Truncated* multivariabel.
2. Mendapatkan estimasi dari kurva regresi Deret Fourier dalam regresi nonparametrik multivariabel.
3. Mengaplikasikan model regresi nonparametrik Spline *Truncated* dan Deret Fourier pada data penduduk miskin di Provinsi Papua.

1.4 Batasan Masalah

Batasan masalah penelitian ini, yaitu:

1. Model regresi nonparametrik yang digunakan adalah Spline *Truncated* dan Deret Fourier.
2. Data yang digunakan adalah data sekunder pada tahun 2012.
3. Metode estimasi yang digunakan dalam penelitian ini adalah *Least Square*.
4. Pemilihan titik knot pada Spline *Truncated* yaitu 3 titik knot.
5. Nilai K pada Deret Fourier ditentukan yaitu K = 1, 2, 3.

1.5 Manfaat Penelitian

1. Memperkaya wawasan dan literature dalam bidang statistika yang berhubungan dengan regresi nonparametrik dengan Spline *Truncated* dan Deret Fourier.
2. Hasil penelitian ini diharapkan dapat memberikan wacana baru bagi bidang ilmu sosial dalam pengambilan keputusan atau kebijakan.
3. Penelitian ini dapat menjadi masukan kepada pihak pemerintahan khususnya pemerintahan Provinsi Papua mengenai permasalahan kependudukan khususnya mengenai pengentasan kemiskinan.

BAB 2
TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan suatu metode Statistika untuk menentukan hubungan antara satu variabel dengan variabel yang lain. Tujuan utama dalam analisis regresi adalah bagaimana mencari bentuk estimasi untuk kurva regresi. Selain untuk mencari bentuk estimasi untuk kurva regresi, analisis regresi juga dapat digunakan untuk prediksi. Misalkan terdapat sekumpulan data berpasangan (x_i , y_i) yang secara umum dapat dimodelkan dengan model regresi:

$$\underset{\sim}{y}_i = \underset{\sim}{f}\left(x_i\right) + \underset{\sim}{\varepsilon}_i ; i = 1, 2, ..., n \tag{2.1}$$

Dengan $\underset{\sim}{y}_i$ respon ke-*i*, $\underset{\sim}{f}\left(x_i\right)$ kurva regresi, $\underset{\sim}{\varepsilon}_i$ *error* yang diasumsikan identik, independen, dan berdistribusi normal.

Berkaitan dengan model tersebut, terdapat tiga model pendekatan regresi, yaitu regresi parametrik, regresi semiparametrik, dan regresi nonparametrik. Apabila dalam analisis regresi bentuk kurva regresi diketahui maka pendekatan model regresi tersebut dinamakan model regresi parametrik (Budiantara, 2007). Sedangkan apabila bentuk kurva regresi tidak diketahui bentuk polanya maka digunakan regresi nonparametrik. Regresi Semiparametrik digunakan jika dalam model regresi yang terdapat komponen Parametrik dan komponen Nonparametrik.

2.2 Regresi Parametrik

Regresi parametrik merupakan metode yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor dengan asumsi bentuk kurva regresi diketahui. Model regresi parametrik dalam bentuk matrik diberikan oleh

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \tag{2.2}$$

dengan:

y : vektor kolom berukuran *nx1* yang elemennya berupa data respon,

X : matrik berukuran *nx(m+1)* dan elemen berupa data *m* prediktor,

$\underset{\sim}{\varepsilon}$: vektor kolom berukuran *nx1* dengan elemen berupa *error random*.

Estimasi parameter pada regresi parametrik diperoleh dengan menggunakan metode *Ordinary Least Square* (OLS) sebagaimana Persamaan (2.3).

$$\hat{\underset{\sim}{\beta}} = \left(X'X\right)^{-1} X'\underset{\sim}{y} \tag{2.3}$$

Estimasi dari y dapat dijabarkan sebagaimana Persamaan (2.4).

$$\begin{aligned}\hat{\underset{\sim}{y}} &= X\hat{\underset{\sim}{\beta}} \\ &= \left(X'X\right)^{-1} X'\underset{\sim}{y}\end{aligned} \tag{2.4}$$

2.3 Regresi Nonparametrik Spline

Spline merupakan bentuk khusus dari potongan (*piecewise*) polynomial orde p yang memiliki sifat tersegmen kontinu sehingga efektif menjelaskan karakteristik lokal dari fungsi data (Eubank, 1988). Spline dalam regresi nonparametrik terus berkembang sampai pada model adaptive (Billier dan Fahrmeir, 2000) dan *multivariate* respon (Holmes dan Mallick, 2003). Metode regresi nonparametrik merupakan metode regresi yang digunakan ketika kurva regresi antara variabel respon dan variabel prediktor tidak diketahui bentuknya atau polanya. Dalam melakukan pemodelan, sebisa mungkin dapat memodelkan secara sederhana. Dalam keadaan dimana terdapat kondisi yang mengharuskan pemodelan yang lebih kompleks, maka model sederhana tidak selayaknya dipaksakan, karena hasil yang diperoleh akan bias dan memiliki *error* yang besar (Budiantara, 2009). Untuk menanggulangi hal ini, perlu dilakukan pendekatan data dengan menggunakan regresi nonparametrik, dimana data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh subjektifitas penelitinya. Dengan kata lain, regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank, 1988).

Model regresi nonparametrik secara umum dapat disajikan sebagai berikut

$$y_i = f\left(x_i\right) + \varepsilon_i, i = 1, 2, ..., n \tag{2.5}$$

dengan $f\left(x_i\right)$ merupakan kurva regresi yang dihampiri dengan fungsi spline berorde p dengan titik knot $K_1, K_2, ..., K_r$ yang dapat diberikan oleh persamaan.

$$f\left(x_i\right)=\sum_{j=0}^{p}\beta_j x_i^j+\sum_{j=0}^{p}\beta_{p+j}\left(x_i-K_j\right)_+^p \tag{2.6}$$

Apabila persamaan (2.5) disubstitusikan kedalam persamaan (2.6) maka akan diperoleh persamaan regresi nonparametrik spline sebagai berikut.

$$y_i=\sum_{j=0}^{p}\beta_j x_i^j+\sum_{j=0}^{p}\beta_{p+j}\left(x_i-K_j\right)_+^p+\varepsilon_i, i=1,2,...,n \tag{2.7}$$

Fungsi $\left(x_i-K_j\right)_+^p$ merupakan fungsi *truncated* (potongan) yang diberikan oleh:

$$\left(x_i-K_j\right)_+^p=\begin{cases}\left(x_i-K_j\right)^p & ,x_i\geq K_j\\ 0 & ,x_i<K_j\end{cases} \tag{2.8}$$

Penelitian estimator spline pada model regresi nonparametrik, telah banyak dilakukan seperti Huang (2003), Crainiceanu, Ruppert dan Wand (2004), Kim dan Gu (2004), Lee (2004), serta Howell (2007). Model regresi nonparametrik yang menyatakan hubungan antara p variabel prediktor dengan variabel respon tunggal dapat ditulis sebagai berikut:

$$y_i=\sum_{j=1}^{p}f_j(x_{ji})+\varepsilon_i,\ i=1,2,...,n \tag{2.9}$$

Persaman (2.9) merupakan model regresi nonparametrik multivariabel yang selanjutnya akan didekati dengan fungsi spline pada persamaan (2.10).

$$f_j\left(x_{ji}\right)=\sum_{v=1}^{m}\beta_{vj}x_{ji}^v+\sum_{k=1}^{r}\beta_{j(m+k)}\left(x_{ji}-K_{jk}\right)_+^m \tag{2.10}$$

dengan,

x_{ji} : variabel prediktor ke j, $j=1, 2, ..., p$ dan $i=1,2,...,n$

y_i : variabel respon ke i, $i=1,2,...,n$

f : kurva regresi

p : banyaknya variabel prediktor

ε_i : error (random) pada subyek ke $i=1,2,...,n$

2.4 Pemilihan Titik Knot Optimal

Dalam spline, titik knot merupakan titik perpaduan yang menunjukkan antara perubahan perilaku fungsi pada interval yang berlainan. Pemilihan titik knot optimal dalam regresi spline berdasarkan pada metode *Generalized Cross Validation* (GCV). Dalam mendapatkan model regresi spline terbaik maka titik optimal dicari yang paling sesuai dengan data. Salah satu metode yang banyak dipakai dalam memilih titik knot optimal adalah *Generalized Cross Validation* (GCV). Jika GCV dibandingkan dengan metode lain, misalnya *Cross Validation* (CV) dan metode *unbiased risk* (UBR)), maka metode ini memiliki sifat optimal asimtotik (Wahba, 1990). Ide dasar dari GCV adalah memodifikasi CV (Green dan Silverman, 1994). Beberapa penulis seperti Craven dan Wahba (1979), Wahba (1985), Li (1986), Kohn (1991), Shao (1993), Venter dan Snyman (1995) mengembangkan metode CV dan GCV. Untuk memperoleh titik knot optimal dapat dilihat dari nilai GCV yang paling minimum. Metode GCV secara umum didefinisikan sebagai berikut (Eubank, 1988).

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1,K_2,\dots,K_r)}{(n^{-1}trace\,[I-A(K_1,K_2,\dots,K_r)])^2} \tag{2.11}$$

dengan

$$MSE(K_1, K_2, \dots, K_r) = n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{f}(x_i)\right)^2 \tag{2.12}$$

2.5 Regresi Nonparametrik Deret Fourier

Estimator Deret Fourier univariabel dalam regresi nonparametrik umumnya digunakan apabila data yang diselidiki polanya tidak diketahui dan ada kecendrungan pola berulang (Tripena dan Budiantara, 2007 dan Bilodeau, 1992). Diberikan model regresi nonparametrik univariabel $y_i = f(x_i) + \varepsilon_i$, $i = 1,2,\dots,n$. Bentuk kurva regresi $f(x)$ diasumsikan tidak diketahui dan termuat di dalam ruang fungsi kontinu $C(0,\pi)$. *Error* random ε_i diasumsikan berdistribusi normal independen dengan mean nol dan variansi σ^2. Karena $f(x)$ kontinu pada interval $(0,\pi)$ maka dapat dihampiri oleh fungsi Deret Fourier $F(x)$, dengan:

$$F(x) = bx + \frac{1}{2}\alpha_0 + \sum_{k=1}^{K} \alpha_k \cos kx \tag{2.13}$$

dimana $b, \alpha_0, \alpha_k,\ k = 1, 2, \ldots, K$ merupakan parameter-parameter model.

Pendekatan regresi nonparametrik Deret Fourier diperoleh dengan meminimumkan *Least Squares* (LS). Misalkan kurva regresi diasumsikan termuat di dalam ruang $C(0,\pi)$:

$$C(0,\pi) = \{f, f \text{ kontinu pada } (0,\pi)\}.$$

Estimator $\hat{f}$ diperoleh dengan meminimumkan *Least Square*:

$$n^{-1}\sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2 \tag{2.14}$$

dengan $\lambda > 0$ adalah parameter penghalus (Bilodeau, 1992).

Estimator Deret Fourier dapat ditulis menjadi :

$$\hat{f}(x_i) = \hat{b}(x_i) + \frac{1}{2}\hat{\alpha}_0 + \sum_{k=1}^{K}\hat{\alpha}_k \cos kx_i, \tag{2.15}$$

(Bilodeau, 1992).

2.6 Beberapa Pengertian Dasar Teorema yang menunjang dalam Aljabar Matriks

Berikut ini diberikan beberapa pengertian dasar yang digunakan dalam mengkaji estimator kurva regresi nonparametrik dengan pendekatan Spline *Truncated* dan Deret Fourier multivariabel:

Teorema 2.6.1 (Searly, 1971)

Diberikan matriks *A* dan *B* yang saling *close form*, maka berlaku sifat – sifat sebagai berikut :

a. Jika matriks *A* simetris, maka *A*=*A*.
b. $(A+B)' = A' + B'$
c. $(AB)' = B'A'$ ∎

Teorema berikutnya berkaitan dengan diferensial pada matriks dan vektor yang diperlukan.

Teorema 2.6.2 (Rencher dan Schaalje, 2008)

Diberikan vektor $\underset{\sim}{a}$ dan $\underset{\sim}{x}$, dimana $\underset{\sim}{a}'\underset{\sim}{x} = \underset{\sim}{x}'\underset{\sim}{a}$, $a' = \left(a_1, a_2, ..., a_p\right)$ memuat konstanta, dengan ini berlaku

$$\frac{\partial\left(\underset{\sim}{a}'\underset{\sim}{x}\right)}{\partial \underset{\sim}{x}} = \frac{\partial\left(\underset{\sim}{x}'\underset{\sim}{a}\right)}{\partial \underset{\sim}{a}} = \underset{\sim}{x} \qquad \blacksquare$$

Teorema 2.6.3 (Rencher dan Schaalje, 2008)

Diberikan vektor $\underset{\sim}{x}$ dan A merupakan suatu matriks simetri, maka

$$\frac{\partial\left(\underset{\sim}{x}'A\underset{\sim}{x}\right)}{\partial \underset{\sim}{x}} = 2A\underset{\sim}{x} \qquad \blacksquare$$

2.7 Pengujian Parameter Model

Pengujian signifikansi parameter pada regresi spline bertujuan untuk mengetahui variabel prediktor berpengaruh nyata atau tidak terhadap variabel respon. Terdapat dua tahap pengujian parameter yaitu pengujian secara serentak dan pengujian secara individu (parsial).

2.7.1 Uji Serentak

Uji serentak yaitu pengujian seluruh parameter yang terdapat dalam model secara bersama- sama. Hipotesis yang digunakan adalah sebagai berikut.

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_m = 0$

H_1 : paling sedikit ada satu $\beta_j \neq 0$, $j = 1,2,\cdots,m$

Dengan statistik uji

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \qquad (2.16)$$

maka daerah penolakannya adalah H_0 ditolak apabila F_{hitung} lebih besar daripada $F_{\alpha(p+r-1,n-p-r)}$. Apabila H_0 ditolak, maka dapat disimpulkan bahwa minimal terdapat satu parameter pada model regresi spline yang signifikan. MS Regresi dan MS Error didapatkan dari Analisis Ragam (ANOVA) sebagaimana yang ditunjukkan Tabel 2.1.

Tabel 2.1 Analisis Ragam (ANOVA)

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F hitung
Regresi	m	$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	$\frac{SS_{regresi}}{df_{regresi}}$	$\frac{MS_{regresi}}{MS_{residual}}$
Residual	$n - m$	$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$\frac{SS_{residual}}{df_{residual}}$	
Total (terkoreksi)	$n - 1$	$\sum_{i=1}^{n}(y_i - \bar{y})^2$	-	

Sumber : (Drapper & Smith, 1992)

Salah satu kriteria untuk mengukur kebaikan model adalah dengan menggunakan R^2. Dimana jika nilai $R^2 > 70\%$ maka dapat dikatakan bahwa model telah dianggap baik (Walpole, 1995). Berikut adalah rumus untuk mendapatkan nilai R^2.

$$R^2 = \frac{SS_{regresi}}{SS_{total}} \tag{2.17}$$

2.7.2 Uji Parsial

Untuk melakukan uji secara parsial pada parameter regresi spline menggunakan uji t. Hipotesis pada uji t adalah sebagai berikut.

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0\ ,\ j = 1,2, \cdots, m$

Statistik uji dari uji parsial adalah sebagai berikut

$$t_{hitung} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \tag{2.18}$$

dengan $se(\hat{\beta}_j)$ adalah *standart error* dari $\hat{\beta}_j$

$$se(\hat{\beta}_j) = \frac{s^2}{\sum_{i=1}^{n}(y_i - \hat{y})^2} \tag{2.19}$$

s^2 merupakan estimator dari varians populasi. H_0 ditolak jika $|t_{hit}| \geq t_{(\frac{\alpha}{2}, n-m)}$.

2.8 Pemeriksaan Asumsi Residual

Residual dari model regresi spline harus memenuhi asumsi IIDN (0, σ^2). Dalam mendeteksi residual tersebut telah memenuhi asumsi maka diperlukan adanya pemeriksaan terhadap residual tersebut.

2.8.1 Uji Identik

Salah satu syarat asumsi residual adalah identik, dimana variansi antar residual harus sama atau tidak terjadi heteroskedastisitas. Tujuan mendeteksi adanya kasus heteroskedastisitas adalah upaya untuk mengurangi kerugian bagi efisiensi estimator (Eubank & Thomas, 1993). Mendeteksi ada tidaknya korelasi dapat dilakukan dengan dua cara yaitu secara visual dengan membuat plot antara $\hat{y}$ dengan residual. Apabila terdapat pola maka dapat diindikasikan terjadi heterokedastisitas dan asumsi identik tidak terpenuhi (Gujarati, 2004).

Selain itu dapat dilakukan dengan uji *Glejser* yang meregresikan harga mutlak residual dengan variabel prediktor (x). Hipotesis yang digunakan adalah sebagai berikut.

$H_0: \sigma_1{}^2 = \sigma_2{}^2 = \cdots = \sigma_n{}^2 = \sigma^2$

$H_1: minimal\ ada\ satu\ \sigma_i{}^2 \neq \sigma^2, i = 1,2,\ldots,n$.

Dimana dan n adalah banyaknya variabel prediktor.

Statistik uji untuk uji *Glejser* adalah.

$$F_{hitung} = \frac{\left[\sum_{i=1}^{n}\left(|\hat{e}_i| - |\bar{e}|\right)^2\right] \Big/ (s-1)}{\left[\sum_{i=1}^{n}\left(|e_i| - |\hat{e}_i|\right)^2\right] \Big/ (n-1)} \tag{2.20}$$

Keputusanya adalah tolak H_0 jika nilai F_{hitung} lebih besar dari $F_{tabel}(F_{\alpha;(s-1,n-s)})$.

2.8.2 Uji Independen

Asumsi residual selanjutnya yang harus terpenuhi independen. Uji independen dilakukan untuk memastikan bahwa tidak terdapat korelasi antar residual atau autokorelasi. Pendeteksian autokorelasi dapat dilakukan dengan melihat plot ACF dari residual. Persamaan kovarian antara Z_t dan Z_{t+k} dapat ditulis sebagai berikut (Wei, 2006).

$$\rho_k = \frac{Cov(Z_t Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\delta_k}{\delta_0} \tag{2.21}$$

dimana:

ρ_k = korelasi antara Z_t dan Z_{t+k}

δ_k = kovarian antara Z_t dan Z_{t+k}

δ_0 = Var (Z_t) = Var (Z_{t+k})

Interval konfidensi dengan batas signifikansi atas dan bawah untuk koefisien ACF diberikan sebagai berikut.

$$0 \pm \frac{Z_{\alpha/2}}{\sqrt{n}} \tag{2.22}$$

2.8.3 Uji Distribusi Normal

Pada model regresi harus memiliki residual yang mengikuti distribusi Normal dengan mean nol dan varians σ^2. Untuk mengetahui apakah residual telah berdistribusi normal dapat dilakukan secara visual dengan *normal probability plot residual.* Selain itu, dapat dilakukan pengujian dengan uji distribusi normal *Kolmogorov-Smirnov*.

Hipotesis untuk uji *Kolmogorov-Smirnov* adalah

$H_0: F_0(x) = F(x)$ (Residual berdistribusi Normal)

$H_1: F_0(x) \neq F(x)$ (Residual tidak berdistribusi Normal)

Statistik uji yang digunakan adalah

$$D = maks|F_0(x) - S_N(x)| \tag{2.23}$$

dimana

$F_0(x)$ = fungsi distribusi frekuensi kumulatif

$S_N(x)$ = k/N adalah fungsi peluang kumulatif yang diobservasi dari suatu sampel random dengan N observasi dimana k adalah banyaknya observasi yang sama atau kurang dari x.

Daerah penolakan H_0 adalah apabila $|D| > q_{(1-\alpha)}$ dengan nilai $q_{(1-\alpha)}$ didapatkan dari tabel *Kolmogorov-Smirnov*.

2.9 Definisi Penduduk Miskin dan Presentase Penduduk Miskin

Negara berkembang dengan jumlah penduduk terbanyak didunia yakni diantaranya adalah India dan Indonesia, tentu memiliki berbagai masalah dalam mengatur kehidupan sosialnya. Jumlah penduduk yang sangat besar dengan karakteristik penduduk yang sebagian besar memiliki tingkat pendidikan yang rendah akan membuat pemerintah sulit mengatur rakyatnya. Masalah yang sering dijumpai pemerintah adalah masalah sosial dari penduduknya. Jumlah penduduk yang banyak dengan sebagian besar penduduknya memiliki tingkat pendidikan yang rendah akan memicu adanya kesenjangan sosial dan terjadi kemiskinan. Salah satu negara yang memiliki masalah dengan persentase kemiskinan adalah Indonesia. Negara Indonesia merupakan negara berkembang dengan jumlah penduduk yang sangat besar. Dengan jumlah penduduk terbanyak keempat didunia dengan sebagian besar penduduknya memiliki tingkat pendidikan rendah, Indonesia tentunya memiliki berbagai masalah dalam mewujudkan kesejahteraan masyarakatnya secara merata. Tujuan dari pembangunan milenium yang telah disepakati oleh anggota Perserikatan Bangsa – Bangsa (PBB) dan Konferensi Tingkat Tinggi adalah *Millenium Development Goals* (MDGs). MDGs yang disepakati sejak tahun 1990 hingga 2015 memiliki tujuan untuk mempercepat pembangunan manusia dan pengentasan kemiskinan. Fokus yang tersirat dari deklarasi ini adalah meningkatkan kesejahteraan manusia dalam berbagai aspek. Salah satu aspek dari kesejahteraan manusia adalah kemiskinan penduduk.

Meskipun Indonesia adalah negara yang kaya akan hasil buminya, namun itu tidak membuat sebagian besar penduduk Indonesia mendapat kehidupan yang layak. Sehingga, kemiskinan terjadi di beberapa daerah yang tersebar di Indonesia. Kemiskinan adalah suatu permasalahan yang memang bersifat kompleks, sehingga diharapkan pemerintah dapat mengentaskan kemiskinan di Indonesia.

Kemiskinan adalah suatu ketidakmampuan untuk memenuhi standar tertentu dari kebutuhan dasar, baik makanan maupun bukan makanan. Suatu ukuran untuk menyatakan standar tersebut adalah berupa garis kemiskinan. Garis kemiskinan merupakan nilai pengeluaran konsumsi kebutuhan dasar makanan setara 2100 kalori energi perkapita perhari, ditambah dengan nilai pengeluaran untuk kebutuhan dasar bukan makanan yang paling pokok.

Menurut Badan Pusat Statistik (BPS) definisi dari penduduk miskin adalah penduduk yang memiliki pendapatan yang berada dibawah garis kemiskinan. Pengertian dari garis kemiskinan sendiri adalah suatu garis yang menunjukkan nilai pengeluaran makanan perorang untuk memenuhi kebutuhan dasar 2100 kkal perhari ditambah dengan pengeluaran non makanan selama 1 bulan. Perhitungan Garis Kemiskinan dilakukan dengan penjumlahan Garis Kemiskinan Makanan dan Garis Kemiskinan Non Makanan.

Penduduk miskin merupakan penduduk yang memiliki pengeluaran perkapita dalam satu bulannya yakni lebih kecil dari garis kemiskinan. Kemiskinan sendiri memiliki arti yakni suatu keadaan dimana terjadi ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan kesehatan. Kemiskinan dapat disebabkan oleh kelangkaan alat pemenuh kebutuhan dasar ataupun sulitnya akses terhadap pendidikan dan pekerjaan.

Persentase penduduk miskin dan pengeluaran perkapita makanan merupakan indikator yang digunakan untuk mengetahui tingkat kesejahteraan penduduk di suatu wilayah. Kajian mengenai kemiskinan rumah tangga pernah dilakukan oleh Chernichovsky dan Meesok (1985) yang mengkaji rumah tangga miskin di Indonesia. Hasil penelitian tersebut menunjukkan bahwa sebagian besar rumah tangga dari penduduk di Indonesia memiliki berbagai karakteristik tersendiri.

Karakteristik rumah tangga miskin tersebut antara lain yakni tingkat pendidikan kepala rumah tangga rendah, pekerjaan yang tidak tetap dari kepala rumah tangga sebagai tulang punggung, sebagian besar pendapatan rumah tangga adalah dari pertanian yang tanahnya dikuasai secara marginal, air bersih yang minim, ketersediaan listrik minim sebagai penerangan serta kondisi rumahnya yang memprihatinkan. Selain itu, presentase kemiskinan serta kedalaman kemiskinan pernah dilakukan oleh banyak peneliti. Berbagai faktor sangat mempengaruhi kemiskinan. Berdasarkan penelitian yang telah dilakukan

Eddy (2010) variabel yang berpengaruh signifikan terhadap kemiskinan pada bidang pendidikan dan jenis pekerjaan. Pada penelitian tersebut dikatakan bahwa pekerjaan di sektor pertanian dan informal memiliki pengaruh yang besar terhadap presentase kemiskinan. Menurut Pintowati (2012), angka melek huruf sangat berpengaruh terhadap presentase kemiskinan. Menurut Eddy (2010) variabel pendidikan kurang dari SD, bekerja di sektor formal dan bekerja di sektor pertanian berpengaruh signifikan terhadap kemiskinan. Di samping itu, penelitian yang telah dilakukan Seruni (2014) variabel pendidikan khususnya rata-rata lama sekolah sangat berpengaruh secara signifikan terhadap kemiskinan.

BAB 3

METODOLOGI PENELITIAN

Untuk menyelesaikan permasalahan dan mencapai tujuan penelitian, dalam bab ini dibahas mengenai langkah-langkah penelitian yang dilaksanakan pada penelitian, aplikasi data dan jadwal kegiatan penelitian yang secara lengkap diuraikan sebagai berikut:

3.1 Mengkaji Estimasi Kurva Regresi Nonparametrik Spline Multivariabel

Mendapatkan estimasi kurva regresi nonparametrik Spline nonparametrik multivariabel langkah-langkahnya adalah sebagai berikut:

a. Membentuk model dengan Nonparametrik Spline Multivariabel

$$y_i = \sum_{j=1}^{p} f_j\left(x_{ji}\right) + \varepsilon_i \; ; \; i = 1,2,\ldots,n$$

b. Menghampiri Kurva Regresi Spline *Truncated* derajat m dengan r titik knot

$$f_j\left(x_{ji}\right) = \sum_{v=1}^{m} \beta_{vj} x_{ji}^{v} + \sum_{k=1}^{r} \beta_{j(m+k)} \left(x_{ji} - K_{jk}\right)_{+}^{m}$$

c. Membentuk optimasi *Least Square*

$$\underset{\underset{\sim}{\beta} \in R^{p(m+r)}}{Min} \left\{ n^{-1} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} \left(\sum_{v=1}^{m} \beta_{vj} x_{ji}^{v} + \sum_{k=1}^{r} \beta_{j(m+k)} \left(x_{ji} - K_{jk}\right)_{+}^{m} \right) \right)^2 \right\}$$

d. Menyajikan persamaan *Least Square* dalam bentuk matriks

$$\underset{\underset{\sim}{\beta} \in R^{p(m+r)}}{Min} \left\{ n^{-1} \left\| \underset{\sim}{y} - X \underset{\sim}{\beta} \right\| \right\}$$

e. Menyelesaikan optimasi (d)dengan derivatif parsial.

3.2 Mengkaji Estimasi Kurva Regresi Deret Fourier dalam Regresi Nonparametrik Multivariabel.

Mendapatkan estimasi kurva regresi Deret Fourier dalam regresi nonparametrik multivariabel, langkah – langkah yang akan dilakukan adalah sebagai berikut:

a. Diberikan model Regresi Nonparametrik Multivariabel

$$y_i = \sum_{j=1}^{q} f_j\left(x_{ji}\right) + \varepsilon_i \; ; \; i = 1, 2, ..., n$$

b. Menghampiri Kurva Regresi dengan Fungsi Deret Fourier

$$f_j\left(x_{ji}\right) = b_j x_{ji} + \frac{1}{2}\alpha_{0j} + \sum_{k=1}^{K} \alpha_{kj} \cos kx_{ji} \; , \; j = 1, 2, ..., q$$

c. Membentuk *Goodness of fit*

$$n^{-1}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{q} f_j\left(x_{ji}\right)\right)^2$$

d. Menyelesaikan optimasi

$$\underset{\alpha \in R^{q(K+2)}}{Min} \left\{ n^{-1}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{q} f_j\left(x_{ji}\right)\right)^2 \right\}$$

3.3 Aplikasi Data untuk Perbandingan Deret Fourier dan *Spline Truncated*

Aplikasi Data yang digunakan dalam penelitian ini yaitu data kemiskinan di Provinsi Papua. Adapun sumber, struktur data dan variabel penelitian adalah sebagai berikut:

3.3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder tahun 2012 yang diperoleh dari publikasi Badan Pusat Statistik (BPS) Provinsi Papua yakni Survey Sosial Ekonomi Nasional (SUSENAS) pada tahun 2012. Unit observasi yang digunakan dalam penelitian ini adalah 29 Kabupaten/Kota yang ada di Provinsi Papua.

Tabel 3.1. Struktur Data Penelitian

Kabupaten/Kota	y_i	x_{1i}	x_{2i}	x_{3i}	x_{4i}	x_{5i}
1	y_1	x_{11}	x_{21}	x_{31}	x_{41}	x_{51}
2	y_2	x_{12}	x_{22}	x_{32}	x_{42}	x_{52}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
29	Y_{29}	x_{129}	x_{229}	x_{329}	x_{429}	X_{529}

3.3.2 Variabel Penelitian

Adapun variabel penelitian yang digunakan yakni terdiri dari variabel respon (*y*) dan variabel prediktor yang ditunjukan oleh Tabel 3.2 berikut ini.

Tabel 3.2. Variabel Penelitian

Variabel	Nama Variabel	Definisi Operasional
y_1	Persentase Penduduk Miskin	Persentase penduduk yang berada dibawah Garis Kemiskinan
x_1	Angka Melek Huruf	Proporsi penduduk berusia 15 tahun ke atas yang dapat membaca dan menulis dalam huruf latin atau lainnya.
x_2	Rata-rata Lama Sekolah	Rata-rata jumlah tahun usia 15 tahun ke atas di seluruh jenjang pendidikan formal yang pernah dijalaninya.
x_3	Pendidikan Kurang dari SD	Presentase penduduk berstatus pendidikan tidak sekolah/belum tamat SD.
x_4	Bekerja di sektor Pertanian	Presentase penduduk yang bekerja di sektor pertanian
x_5	Bekerja di sektor informal	Presentase penduduk yang bekerja di sektor informal

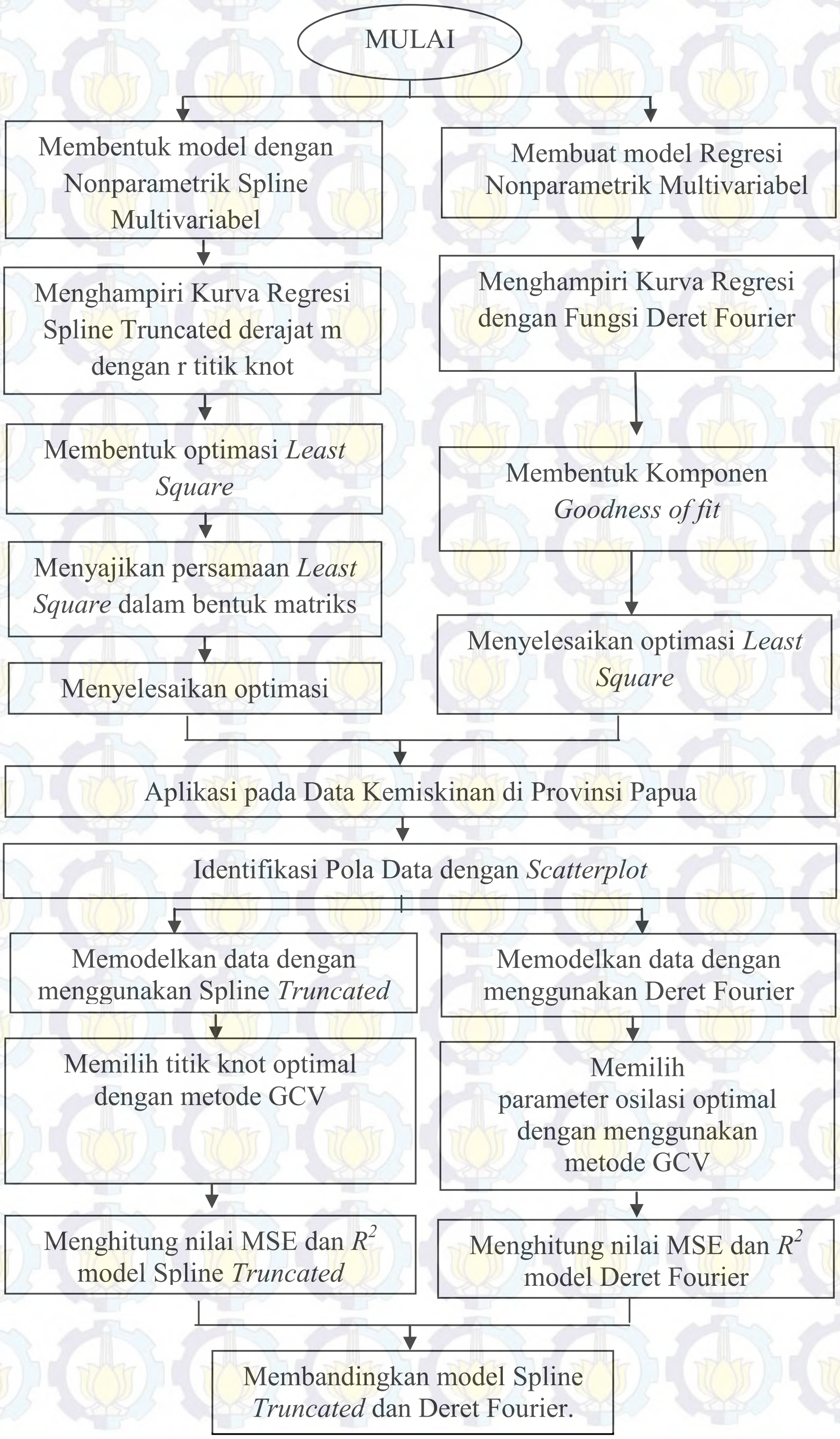
Pemilihan variabel prediktor yang digunakan berdasarkan penelitian yang telah dilakukan Eddy (2010) variabel yang berpengaruh signifikan terhadap kemiskinan pada bidang Pendidikan dan Jenis Pekerjaan. Pada penelitian tersebut dikatakan bahwa pekerjaan di sektor pertanian dan informal memiliki pengaruh yang besar terhadap presentase kemiskinan. Menurut Pintowati (2012), Angka Melek Huruf sangat berpengaruh terhadap presentase kemiskinan. Menurut Eddy (2010) variabel pendidikan kurang dari SD, bekerja di sektor formal dan bekerja di sektor pertanian berpengaruh signifikan terhadap kemiskinan. Di samping itu, penelitian yang telah dilakukan Seruni (2014) variabel pendidikan khususnya rata-rata lama sekolah sangat berpengaruh secara signifikan terhadap kemiskinan.

3.3.3 Aplikasi model pada Data Kemiskinan di Provinsi Papua

Langkah-langkah untuk aplikasi model data kemiskinan di Provinsi Papua adalah sebagai berikut:

a. Membuat *Scatterplot* $(x_{1i}, y_i), (x_{2i}, y_i), (x_{3i}, y_i), (x_{4i}, y_i), (x_{5i}, y_i)$;
b. Memodelkan data dengan menggunakan Spline *Truncated*.
c. Memilih titik knot optimal dengan metode GCV.
d. Menghitung nilai MSE dan R^2 model Spline *Truncated* yang terbaik.
e. Memodelkan data dengan menggunakan Deret Fourier.
f. Memilih parameter osilasi optimal dengan menggunakan metode GCV.
g. Menghitung nilai MSE dan R^2 model Deret Fourier yang terbaik.
h. Membandingkan model Spline *Truncated* dan Deret Fourier.

Adapun langkah-langkah penelitiannya dapat digambarkan sebagai berikut



Gambar 3.1 Langkah-langkah Penelitian

(...Halaman Ini Sengaja Dikosongkan...)

BAB 4
HASIL DAN PEMBAHASAN

Berdasarkan metode penelitian yang digunakan, bab ini membahas mengenai hasil dan analisis yang telah dilakukan sesuai dengan tujuan dari penelitian yang tertulis pada Bab 1, yaitu mengkaji cara mendapatkan estimasi kurva regresi Spline *Truncated*, mendapatkan estimasi kurva regresi Deret Fourier dalam regresi nonparametrik multivariabel dan terapannya untuk membandingkan antara Spline *Truncated* dan Deret Fourier pada pemodelan data kemiskinan di Provinsi Papua.

4.1 Estimasi Kurva Regresi Nonparametrik Spline Multivariabel

Dalam bagian ini dibahas tentang estimasi Spline multivariabel. Spline merupakan jumlahan dari fungsi polinomial dengan suatu fungsi (*truncated*). Diberikan model regresi nonparametrik multivariabel:

$$
\begin{aligned}
y_i &= \mu\left(x_{1i}, x_{2i}, \ldots, x_{pi}\right) + \varepsilon_i \\
&= \sum_{j=1}^{p} f_j\left(x_{ji}\right) + \varepsilon_i ; i = 1, 2, \ldots, n
\end{aligned}
$$

Selanjutnya, kurva regresi f_j dihampiri dengan fungsi Spline multivariabel maka dapat ditulis menjadi:

$$
\mu\left(x_{1i}, x_{2i}, \ldots, x_{pi}\right) = \sum_{j=1}^{p} f_j\left(x_{ji}\right) + \varepsilon_i ; i = 1, 2, \ldots, n
$$

dengan:

$$
\begin{aligned}
f_j\left(x_{ji}\right) &= \sum_{v=1}^{m} \beta_{vj} x_{ji}^{v} + \sum_{k=1}^{r} \beta_{j(k+m)} \left(x_{ji} - K_{jk}\right)_+^m \\
&= \beta_{1j} x_{ji}^{1} + \ldots + \beta_{mj} x_{ji}^{m} + \beta_{j(1+m)} \left(x_{ji} - K_{j1}\right)_+^m + \ldots + \beta_{j(r+m)} \left(x_{ji} - K_{jr}\right)_+^m
\end{aligned}
$$

dimana $j = 1, 2, \ldots, p$; $v = 1, 2, \ldots, m$; dan $k = 1, 2, \ldots, r$ titik-titik knot yang memperlihatkan perubahan pola perilaku dari fungsi tersebut pada sub-sub interval yang berbeda. Dengan demikian model regresi dapat ditulis menjadi:

$$\sum_{j=1}^{p} f_j\left(x_{ji}\right)=\sum_{j=1}^{p}\left(\beta_{1j}x_{ji}^{1}+\ldots+\beta_{mj}x_{ji}^{m}+\beta_{j(1+m)}\left(x_{ji}-K_{j1}\right)_{+}^{m}+\ldots+\beta_{j(r+m)}\left(x_{ji}-K_{jr}\right)_{+}^{m}\right)$$

$$=\left(\beta_{11}x_{1i}^{1}+\ldots+\beta_{m1}x_{1i}^{m}+\beta_{1(1+m)}\left(x_{1i}-K_{11}\right)_{+}^{m}+\ldots+\beta_{1(r+m)}\left(x_{1i}-K_{1r}\right)_{+}^{m}\right)+\ldots+$$

$$\left(\beta_{1p}x_{pi}^{1}+\ldots+\beta_{mp}x_{pi}^{m}+\beta_{p(1+m)}\left(x_{pi}-K_{p1}\right)_{+}^{m}+\ldots+\beta_{p(r+m)}\left(x_{pi}-K_{pr}\right)_{+}^{m}\right) \quad (4.1)$$

Apabila persamaan (4.1) disajikan dalam bentuk matrik, maka didapat:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}=\begin{bmatrix} x_{11}^{1} & \cdots & x_{11}^{m} & \left(x_{11}-K_{11}\right)_{+}^{m} & \cdots & \left(x_{11}-K_{1r}\right)_{+}^{m} \\ x_{12}^{1} & \cdots & x_{12}^{m} & \left(x_{12}-K_{11}\right)_{+}^{m} & \cdots & \left(x_{12}-K_{1r}\right)_{+}^{m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{1n}^{1} & \cdots & x_{1n}^{m} & \left(x_{1n}-K_{11}\right)_{+}^{m} & \cdots & \left(x_{1n}-K_{1r}\right)_{+}^{m} \end{bmatrix}\begin{bmatrix} \beta_{11} \\ \vdots \\ \beta_{m1} \\ \beta_{1(1+m)} \\ \vdots \\ \beta_{1(r+m)} \end{bmatrix}+\ldots+$$

$$\begin{bmatrix} x_{p1}^{1} & \cdots & x_{p1}^{m} & \left(x_{p1}-K_{p1}\right)_{+}^{m} & \cdots & \left(x_{p1}-K_{pr}\right)_{+}^{m} \\ x_{p2}^{1} & \cdots & x_{p2}^{m} & \left(x_{p2}-K_{p1}\right)_{+}^{m} & \cdots & \left(x_{p2}-K_{pr}\right)_{+}^{m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{pn}^{1} & \cdots & x_{pn}^{m} & \left(x_{pn}-K_{p1}\right)_{+}^{m} & \cdots & \left(x_{pn}-K_{pr}\right)_{+}^{m} \end{bmatrix}\begin{bmatrix} \beta_{1p} \\ \vdots \\ \beta_{mp} \\ \beta_{p(1+m)} \\ \vdots \\ \beta_{p(r+m)} \end{bmatrix}+\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (4.2)$$

Persamaan (4.2) dapat ditulis menjadi:

$$\underset{\sim}{y}=\left[\begin{matrix} A_1 & \vdots & \cdots & \vdots & A_p \end{matrix}\right]\underset{\sim}{\beta}+\underset{\sim}{\varepsilon}$$

dengan,

$$\underset{\sim}{y}=\left[y_1,\ldots,y_n\right]'$$

$$A_1=\begin{bmatrix} x_{11}^{1} & \cdots & x_{11}^{m} & \left(x_{11}-K_{11}\right)_{+}^{m} & \cdots & \left(x_{11}-K_{1r}\right)_{+}^{m} \\ x_{12}^{1} & \cdots & x_{12}^{m} & \left(x_{12}-K_{11}\right)_{+}^{m} & \cdots & \left(x_{12}-K_{1r}\right)_{+}^{m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{1n}^{1} & \cdots & x_{1n}^{m} & \left(x_{1n}-K_{11}\right)_{+}^{m} & \cdots & \left(x_{1n}-K_{1r}\right)_{+}^{m} \end{bmatrix}$$

$$A_p = \begin{bmatrix} x_{p1}^1 & \cdots & x_{p1}^m & \left(x_{p1}-K_{p1}\right)_+^m & \cdots & \left(x_{p1}-K_{pr}\right)_+^m \\ x_{p2}^1 & \cdots & x_{p2}^m & \left(x_{p2}-K_{p1}\right)_+^m & \cdots & \left(x_{p2}-K_{pr}\right)_+^m \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{pn}^1 & \cdots & x_{pn}^m & \left(x_{pn}-K_{p1}\right)_+^m & \cdots & \left(x_{pn}-K_{pr}\right)_+^m \end{bmatrix}$$

$$\underset{\sim}{\beta} = \left[\beta_{11} \quad \cdots \quad \beta_{m1} \quad \beta_{1(1+m)} \quad \cdots \quad \beta_{1(r+m)} \quad \vdots \cdots \vdots \quad \beta_{1p} \quad \cdots \quad \beta_{mp} \quad \beta_{p(1+m)} \quad \cdots \quad \boldsymbol{\beta_{p(r+m)}}\right]'$$

dan $\underset{\sim}{\varepsilon} = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]'$.

Atau dapat ditulis menjadi:

$$\underset{\sim}{y} = X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

dengan,

$$\underset{\sim}{y} = [y_1,...,y_n]'$$

$$X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) = \left[A_1 \quad \vdots \quad \cdots \quad \vdots \quad A_p\right]$$

$$\underset{\sim}{\beta} = \left(\underset{\sim}{\beta}_1',...,\underset{\sim}{\beta}_p'\right)'$$

$$\underset{\sim}{\beta}_1 = \left(\beta_{11},...,\beta_{m1},\beta_{1(1+m)},...,\beta_{1(r+m)}\right)',...,\underset{\sim}{\beta}_p = \left(\beta_{1p},...,\beta_{mp},\beta_{p(1+m)},...,\beta_{p(r+m)}\right)'$$

dan $\underset{\sim}{\varepsilon} = (\varepsilon_1,...,\varepsilon_n)'$

Estimasi untuk $\underset{\sim}{\beta} = \left(\underset{\sim}{\beta}_1,...,\underset{\sim}{\beta}_p\right)'$ diperoleh dengan menggunakan metode *Least Square*. Estimator parameter $\underset{\sim}{\beta}$ didapat dari menyelesaikan optimasi:

$$\underset{\underset{\sim}{\beta}\varepsilon R^{p(m+r)}}{Min}\left\{n^{-1}\left\|\underset{\sim}{y} - X\underset{\sim}{\beta}\right\|^2\right\}$$

$$= \left\{\left(\underset{\sim}{y} - X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta}\right)'\left(\underset{\sim}{y} - X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta}\right)\right\}$$

Untuk menyelesaikan optimasi dengan menggunakan derivatif parsial, misalkan:

$$\mathbb{Q}\left(\underset{\sim}{\beta}\right) = \left\{\left(\underset{\sim}{y} - X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta}\right)'\right.$$

$$\left.\left(\underset{\sim}{y} - X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta}\right)\right\} \tag{4.3}$$

Berdasarkan Teorema 2.6.1, persamaan (4.3) menjadi:

$$
\begin{aligned}
&= \left(\underset{\sim}{y}' - \underset{\sim}{\beta}' X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\right)\left(\underset{\sim}{y} - X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta}\right) \\
&= \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\beta}' X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y} - \left(\underset{\sim}{y}' X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta}\right) \\
&\quad + \underset{\sim}{\beta}' X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta} \\
&= \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\beta}' X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y} - \left(\underset{\sim}{\beta}' X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y}\right)' \\
&\quad + \underset{\sim}{\beta}' X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta}' \\
&= \underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{\beta} X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y} \\
&\quad + \underset{\sim}{\beta}' X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta} \quad (4.4)
\end{aligned}
$$

Menggunakan dasar Teorema 2.6.2 dan 2.6.3, langkah berikutnya persamaan (4.4) diturunkan terhadap $\underset{\sim}{\beta}$:

$$
\begin{aligned}
\frac{\partial \mathbb{Q}\left(\underset{\sim}{\beta}\right)}{\partial \underset{\sim}{\beta}} &= -2X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y} \\
&\quad + 2X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{\beta} \quad (4.5)
\end{aligned}
$$

Setelah persamaan (4.5) diturunkan hasilnya disamakan dengan 0, maka diperoleh persamaan:

$$
\begin{aligned}
0 &= -2X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y} + \\
&\quad 2X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\hat{\underset{\sim}{\beta}} \quad (4.6)
\end{aligned}
$$

Hasil yang diperoleh dari persamaaan (4.6) adalah

$$
X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\hat{\underset{\sim}{\beta}} = X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y}
$$

Sehingga estimator $\hat{\underset{\sim}{\beta}}$ diberikan oleh:

$$
\hat{\underset{\sim}{\beta}} = \left(X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right) X\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\right)^{-1} X'\left(K_{11},...,K_{1r} \vdots \cdots \vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y}
$$

dengan $\hat{\underset{\sim}{\beta}} = \left(\hat{\underset{\sim}{\beta}}_1',...,\hat{\underset{\sim}{\beta}}_p'\right)'$.

Estimator kurva regresi $\hat{\underset{\sim}{f}}(x)$ diperoleh dari:

$$\hat{\underset{\sim}{f}}(x)=\left[\hat{\underset{\sim}{f}}(x_1)',\hat{\underset{\sim}{f}}(x_2)',...,\hat{\underset{\sim}{f}}(x_p)'\right]'$$

$$=X\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)\hat{\underset{\sim}{\beta}}$$

$$=\left(X'\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)X\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)\right)^{-1}$$

$$X'\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y}$$

$$=\mathrm{B}\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y} \tag{4.7}$$

Matrik $\mathrm{B}\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)$ pada persamaan (4.7) adalah matrik yang diberikan oleh:

$$\left(X'\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)X\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)\right)^{-1}X'\left(K_{11},...,K_{1r}\vdots\cdots\vdots K_{p1},...,K_{pr}\right)\underset{\sim}{y}$$

Estimasi f dapat ditulis menjadi:

$$\hat{f}_j\left(x_{ji}\right)=\sum_{v=1}^{m}\hat{\beta}_{vj}x_{ji}^{v}+\sum_{k=1}^{r}\hat{\beta}_{j(k+m)}\left(x_{ji}-K_{jk}\right)_{+}^{m}$$

$$\sum_{j=1}^{p}\hat{f}_j\left(x_{ji}\right)=\sum_{j=1}^{p}\left(\sum_{v=1}^{m}\hat{\beta}_{vj}x_{ji}^{v}+\sum_{k=1}^{r}\hat{\beta}_{j(k+m)}\left(x_{ji}-K_{jk}\right)_{+}^{m}\right)$$

Akibatnya estimasi untuk kurva regresi $\mu\left(x_{1i},x_{2i},...,x_{pi}\right)$ diberikan oleh:

$$\hat{\mu}\left(x_{1i},x_{2i},...,x_{pi}\right)=\sum_{j=1}^{p}\hat{f}_j\left(x_{ji}\right)$$

$$=\sum_{j=1}^{p}\left(\sum_{v=1}^{m}\hat{\beta}_{vj}x_{ji}^{v}+\sum_{k=1}^{r}\hat{\beta}_{j(k+m)}\left(x_{ji}-K_{jk}\right)_{+}^{m}\right).$$

$$=\sum_{j=1}^{p}\sum_{v=1}^{m}\hat{\beta}_{vj}x_{ji}^{v}+\sum_{j=1}^{p}\sum_{k=1}^{r}\hat{\beta}_{j(k+m)}\left(x_{ji}-K_{jk}\right)_{+}^{m}$$

dimana $\hat{\beta}_{vj}$ dan $\hat{\beta}_{j(k+m)}$ $,v=1,2,...,m;k=1,2,...,r,$ dan $j=1,2,...,p$

Diperoleh dari:

$\hat{\underset{\sim}{\beta}}=\left(\hat{\underset{\sim}{\beta}}_1',...,\hat{\underset{\sim}{\beta}}_p'\right)'$ dan

$$\hat{\underset{\sim}{\beta}}_1=\left(\beta_{11},...,\beta_{m1},\beta_{1(1+m)},...,\beta_{1(r+m)}\right)',...,\hat{\underset{\sim}{\beta}}_p=\left(\beta_{1p},...,\beta_{mp},\beta_{p(1+m)},...,\beta_{p(r+m)}\right)'$$

4.2 Estimasi Kurva Regresi Deret Fourier dalam Regresi Nonparametrik Multivariabel

Pada bagian ini disajikan kajian estimasi Deret Fourier dalam regresi nonparametrik multivariabel. Deret Fourier merupakan polinomial trigonometri yang mempunyai fleksibelitas yang tinggi. Deret Fourier baik digunakan untuk mengetimasi kurva regresi yang menunjukkan gelombang sinus atau cosinus. Diberikan model regresi nonparametrik multivariabel:

$$\begin{aligned} y_i &= \mu\left(x_{1i}, x_{2i}, \ldots, x_{qi}\right) + \varepsilon_i \\ &= \sum_{j=1}^{q} f_j\left(x_{ji}\right) + \varepsilon_i,\ \ i = 1, 2, \ldots, n \end{aligned}$$

Selanjutnya, kurva regresi f_j dihampiri dengan fungsi Deret Fourier:

$$f_j\left(x_{ji}\right) = b_j x_{ji} + \frac{1}{2}\alpha_{0j} + \sum_{k=1}^{K} \alpha_{kj} \cos kx_{ji},\ j = 1, 2, \ldots, q.$$

Estimator kurva regresi Deret Fourier diperoleh dari optimasi:

$$\underset{\underset{\sim}{\beta} \in R^{q(K+2)}}{Min} \left\{ \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{q} \left(b_j x_{ji} + \frac{1}{2}\alpha_{0j} + \sum_{k=1}^{K} \alpha_{kj} \cos kx_{ji} \right) \right)^2 \right\} = \underset{\underset{\sim}{\beta} \in R^{q(K+2)}}{Min} \left\{ \mathbb{Q}\left(\underset{\sim}{\beta}\right) \right\}$$

Untuk menyelesaikan optimasi diatas dilakukan langkah-langkah sebagai berikut:

$$\begin{aligned} \mathbb{Q}\left(\underset{\sim}{\beta}\right) &= \sum_{i=1}^{n} \left(y_i - \mu\left(x_{1i}, x_{2i}, \ldots, x_{qi}\right) \right) \\ &= \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{q} \left(b_j x_{ji} + \frac{1}{2}\alpha_{0j} + \alpha_{1j} \cos x_{ji} + \ldots + \alpha_{Kj} \cos Kx_{ji} \right) \right)^2 \\ &= \sum_{i=1}^{n} \left(y_i - \left(b_1 x_{1i} + \frac{1}{2}\alpha_{01} + \alpha_{11} \cos x_{1i} + \ldots + \alpha_{K1} \cos Kx_{1i} \right) - \ldots + \right. \\ &\quad \left. - \left(b_q x_{qi} + \frac{1}{2}\alpha_{0q} + \alpha_{1q} \cos x_{qi} + \ldots + \alpha_{Kq} \cos Kx_{qi} \right) \right)^2 \end{aligned} \tag{4.8}$$

Persamaan (4.8) diatas dapat ditulis menjadi:

$$\mathbb{Q}\left(\underset{\sim}{\beta}\right) = \left(\underset{\sim}{y} - X(K)\underset{\sim}{\beta}\right)' \left(\underset{\sim}{y} - X(K)\underset{\sim}{\beta}\right) \tag{4.9}$$

dengan:

$\underset{\sim}{y} = [y_1, y_2, ..., y_n]'$,

$\underset{\sim}{\beta} = \left[b_1 \quad \frac{1}{2}\alpha_{01} \quad \alpha_{11} \quad \cdots \quad \alpha_{K1} \quad \vdots \quad \cdots \quad \vdots \quad b_q \quad \frac{1}{2}\alpha_{0q} \quad \alpha_{1q} \quad \cdots \quad \alpha_{Kq} \right]'$ dan,

$$X(K) = \begin{bmatrix} x_{11} & 1 & \cos x_{11} & \cdots & \cos Kx_{11} & \vdots & \cdots & \vdots & x_{q1} & 1 & \cos x_{q1} & \cdots & \cos Kx_{q1} \\ x_{12} & 1 & \cos x_{12} & \cdots & \cos Kx_{12} & \vdots & \cdots & \vdots & x_{q2} & 1 & \cos x_{q2} & \cdots & \cos Kx_{q2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1n} & 1 & \cos x_{1n} & \cdots & \cos Kx_{1n} & \vdots & \cdots & \vdots & x_{qn} & 1 & \cos x_{qn} & \cdots & \cos Kx_{qn} \end{bmatrix}$$

Berdasarkan Teorema 2.6.1, persamaan (4.9) dapat ditulis:

$$\begin{aligned} \mathbb{Q}(\underset{\sim}{\beta}) &= \left(\underset{\sim}{y}' - \underset{\sim}{\beta}'X'(K)\right)\left(\underset{\sim}{y} - X(K)\underset{\sim}{\beta}\right) \\ &= \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{y}'X(K)\underset{\sim}{\beta} - \underset{\sim}{\beta}'X'(K)\underset{\sim}{y} + \underset{\sim}{\beta}'X'(K)X(K)\underset{\sim}{\beta} \\ &= \underset{\sim}{y}'\underset{\sim}{y} - \left(X'(K)\underset{\sim}{\beta}'\underset{\sim}{y}\right)' - \underset{\sim}{\beta}'X'(K)\underset{\sim}{y} + \underset{\sim}{\beta}'X'(K)X(K)\underset{\sim}{\beta} \\ &= \underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{\beta}'X'(K)\underset{\sim}{y} + \underset{\sim}{\beta}'X'(K)X(K)\underset{\sim}{\beta} \end{aligned}$$

Estimasi kurva regresi $\hat{\mu}$ diperoleh dengan cara ekuivalen dengan estimasi $\hat{\underset{\sim}{\beta}}$. Estimasi $\hat{\underset{\sim}{\beta}}$ diperoleh dari meminimumkan $\mathbb{Q}(\underset{\sim}{\beta})$. Menggunakan dasar Teorema 2.6.2 dan 2.6.3 dengan menurunkan secara parsial $\mathbb{Q}(\underset{\sim}{\beta})$ terhadap $\underset{\sim}{\beta}$ didapat:

$$\frac{\partial \mathbb{Q}(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}} = -2X'(K)\underset{\sim}{y} + 2X'(K)X(K)\underset{\sim}{\beta} \tag{4.10}$$

Jika persamaan (4.10) disamakan dengan nol diperoleh persamaan:

$$-2X'(K)\underset{\sim}{y} + 2X'(K)X(K)\hat{\underset{\sim}{\beta}} = 0.$$

Dengan sedikit penjabaran dan menganggap matriks $X(K)$ nonsingular (matriks dengan rank penuh) maka diperoleh estimator $\underset{\sim}{\beta}$ sebagai berikut:

$$\begin{aligned} \hat{\underset{\sim}{\beta}}(K) &= \left(X'(K)X(K)\right)^{-1} X'(K)\underset{\sim}{y} \\ &= \left[\hat{b}_1 \quad \frac{1}{2}\hat{\alpha}_{01} \quad \hat{\alpha}_{11} \quad \cdots \quad \hat{\alpha}_{K1} \quad \vdots \quad \cdots \quad \vdots \quad \hat{b}_q \quad \frac{1}{2}\hat{\alpha}_{0q} \quad \hat{\alpha}_{1q} \quad \cdots \quad \hat{\alpha}_{Kq} \right]' \end{aligned}$$

Estimator untuk kurva regresi f_j diberikan oleh:

$$\hat{f}_j\left(x_{ji}\right) = \hat{b}_j x_{ji} + \frac{1}{2}\hat{\alpha}_{0j} + \sum_{k=1}^{K}\hat{\alpha}_{kj}\cos kx_{ji}.$$

Akibatnya estimasi untuk kurva regresi $\mu\left(x_{1i}, x_{2i}, \ldots, x_{qi}\right)$ diberikan oleh:

$$\hat{\mu}\left(x_{1i}, x_{2i}, \ldots, x_{qi}\right) = \sum_{j=1}^{q}\hat{f}_j\left(x_{ji}\right)$$

$$= \sum_{j=1}^{q}\left(\hat{b}_j x_{ji} + \frac{1}{2}\hat{\alpha}_{0j} + \sum_{k=1}^{K}\hat{\alpha}_{kj}\cos kx_{ji}\right).$$

4.3 Aplikasi Data

Data yang digunakan adalah data kemiskinan di Provinsi Papua, dengan banyak data n = 29, variabel respon adalah presentase kemiskinan. Penelitian ini adalah mengaplikasikan Deret Fourier dan Spline *Truncated* yang selanjutnya menbandingkan kedua metode untuk memodelkan presentase kemiskinan di Provinsi Papua, dengan variabel-variabel yang mempengaruhi.

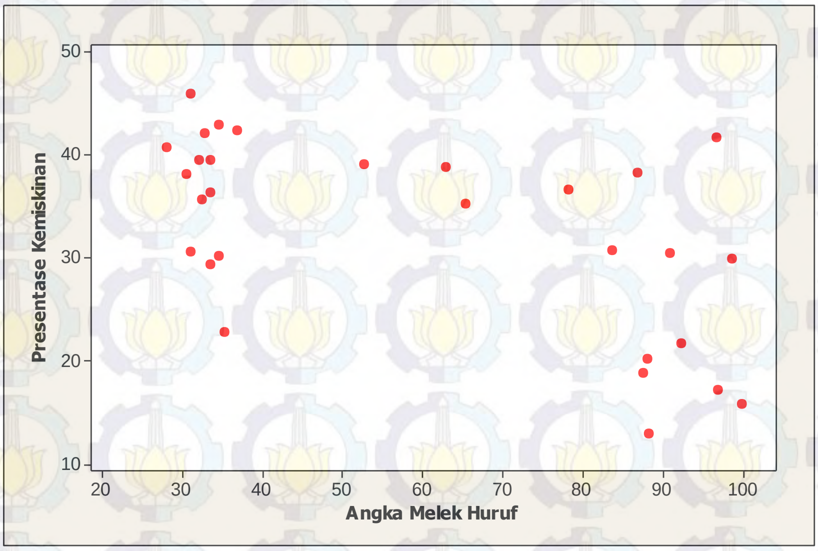
4.3.1 Hubungan Penduduk Miskin dengan Faktor-Faktor yang Diduga Mempengaruhi

Dalam melakukan pemodelan penduduk miskin harus diketahui pola data dari kedua variabel tersebut. Pola dari suatu data dapat berbentuk linear, kuadratik, kubik dan lain sebagainya. Data yang memiliki suatu pola hubungan antara variabel respon dan variabel prediktornya adalah diketahui bentuknya maka analisis dari data tersebut dapat dilakukan menggunakan model regresi dengan pendekatan parametrik. Sebaliknya, apabila pola hubungan antara variabel respon dan variabel prediktor tidak dapat diketahui maka pendekatan yang dilakukan adalah pendekatan regresi nonparametrik. Namun, jika pola hubungan data antara variabel respon dan variabel prediktor sebagian diketahui dan sebagian lagi tidak, maka dilakukan pendekatan regresi semiparametrik.

Berikut akan dilakukan analisa terhadap pola data antara variabel respon presentase penduduk miskin dengan masing–masing variabel–variabel prediktor yakni angka melek huruf, rata-rata lama sekolah, pendidikan kurang dari SD, bekerja di sektor pertanian dan bekerja di sektor informal.

4.3.2 *Scatterplot* Persentase Penduduk Miskin dengan Angka Melek Huruf

Identifikasi menggunakan *scatterplot* akan dapat diketahui pola hubungan antara variabel respon dan variabel prediktor. Faktor pertama yang diduga mempengaruhi variabel respon yaitu presentase penduduk miskin adalah angka melek huruf. Berikut ini adalah gambar yang menunjukkan pola hubungan dari variabel respon dengan variabel angka melek huruf.

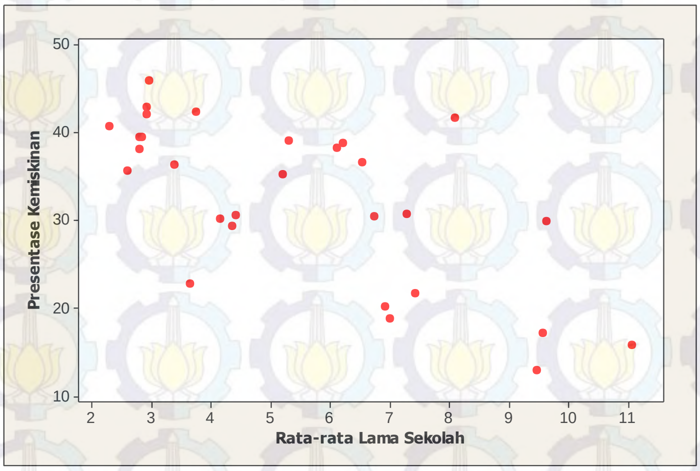


Gambar 4.1 *Scatterplot* Variabel Respon y_1 dengan Variabel Prediktor x_1

Berdasarkan Gambar 4.1 terlihat bahwa pola hubungan variabel respon (presentase penduduk miskin) dengan variabel prediktor (angka melek huruf) menunjukkan pola hubungan yang tidak mengikuti suatu pola tertentu. Pola hubungan yang tidak diketahui polanya inilah menjadikan penyelesaian masalah ini dengan pendekatan regresi nonparametrik.

4.3.3 *Scatterplot* Persentase Penduduk Miskin dengan Rata-rata Lama Sekolah

Faktor kedua yang diduga mempengaruhi variabel respon adalah rata-rata lama sekolah. Berikut ini adalah gambar yang menunjukkan pola hubungan dari variabel respon dengan variabel rata-rata lama sekolah.

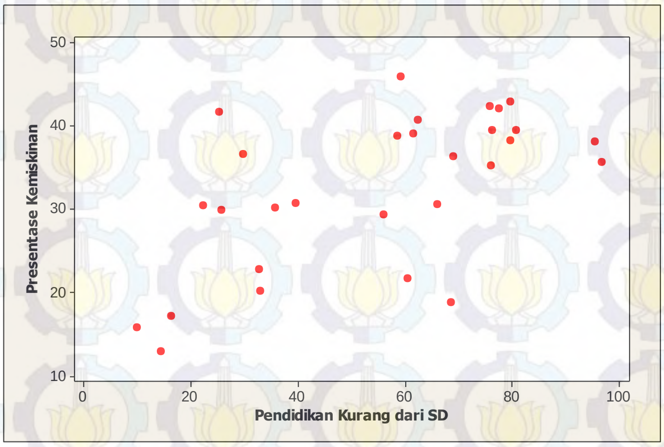


Gambar 4.2 *Scatterplot* Variabel Respon y_1 dengan Variabel Prediktor x_2

Berdasarkan Gambar 4.2 terlihat bahwa pola hubungan variabel respon (presentase penduduk miskin) dengan variabel prediktor (rata-rata lama sekolah) menunjukkan pola hubungan yang tidak mengikuti suatu pola tertentu. Sehingga penyelesaian masalah ini dapat menggunakan pendekatan regresi nonparametrik.

4.3.4 *Scatterplot* Persentase Penduduk Miskin dengan Pendidikan kurang dari Sekolah Dasar (SD)

Faktor ketiga yang diduga mempengaruhi variabel respon adalah pendidikan kurang dari SD. Berikut ini adalah gambar yang menunjukkan pola hubungan dari variabel respon dengan variabel pendidikan kurang dari SD.

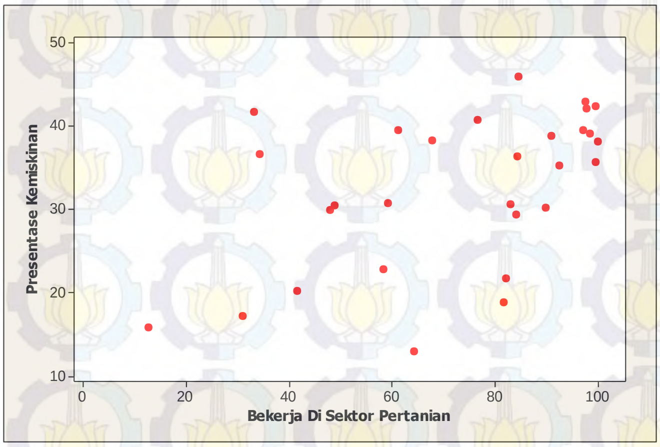


Gambar 4.3 *Scatterplot* Variabel Respon y_1 dengan Variabel Prediktor x_3

Pada Gambar 4.3 terlihat bahwa pola hubungan variabel respon yaitu presentase penduduk miskin dengan variabel prediktor (pendidikan kurang dari SD) menunjukkan pola hubungan yang tidak mengikuti suatu pola tertentu. Sehingga, untuk menyelesaikan permasalahan ini menggunakan pendekatan regresi nonparametrik.

4.3.5 *Scatterplot* Persentase Penduduk Miskin dengan Bekerja di Sektor Pertanian

Variabel keempat yang diduga mempengaruhi variabel respon presentase penduduk miskin adalah bekerja di sektor pertanian. Berikut ini adalah gambar yang menunjukkan pola hubungan dari variabel respon dengan variabel bekerja di sektor pertanian.

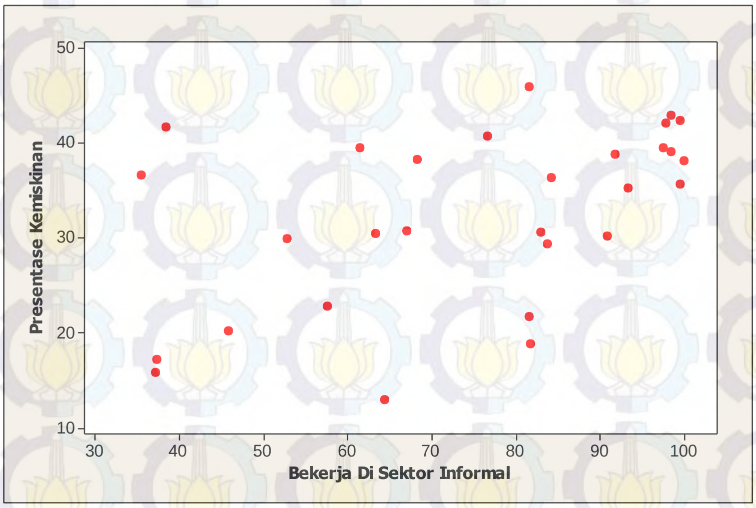


Gambar 4.4 *Scatterplot* Variabel Respon y_1 dengan Variabel Prediktor x_4

Berdasarkan Gambar 4.4 terlihat bahwa pola hubungan variabel respon dengan variabel prediktor (bekerja di sektor pertanian) menunjukkan pola hubungan yang tidak mengikuti suatu pola tertentu. Selanjutnya, dari pola hubungan yang tidak diketahui tersebut maka untuk menyelesaikan masalah ini dilakukan pendekatan regresi nonparametrik.

4.3.6 *Scatterplot* Persentase Penduduk Miskin dengan Bekerja di Sektor Informal

Faktor kelima yang diduga mempengaruhi variabel respon adalah bekerja di sektor informal. Berikut ini adalah gambar yang menunjukkan pola hubungan dari variabel respon dengan variabel bekerja di sektor informal.



Gambar 4.5 *Scatterplot* Variabel Respon y_1 dengan Variabel Prediktor x_5

Pada Gambar 4.5 terlihat bahwa pola hubungan variabel respon dengan variabel bekerja di sektor informal menunjukkan pola hubungan yang tidak mengikuti suatu pola tertentu. Penyelesaian masalah ini dengan pendekatan nonparametrik. Kelima faktor yang diduga mempengaruhi variabel respon tersebut, memiliki pola data yang tidak diketahui polanya sehingga penyelesaiannya menggunakan pendekatan regresi nonparametrik. Pada penelitian ini akan digunakan pendekatan nonparametrik menggunakan regresi Spline *Truncated* dan Deret Fourier.

4.4 Model Regresi Nonparametrik Spline *Truncated*

Pada pembahasan ini akan dilakukan pemodelan mengenai variabel-variabel yang diduga mempengaruhi kemiskinan. Metode yang digunakan untuk memodelkan data ini adalah regresi nonparametrik Spline *Truncated*. Alasan penggunaan metode ini telah dijelaskan pada bagian sebelumnya. Berikut adalah model matematis dari regresi nonparametrik Spline *Truncated*.

$$y_i = \sum_{j=1}^{p}\left(\sum_{v=1}^{m} \hat{\beta}_{vj} x_{ji}^{v} + \sum_{k=1}^{r} \hat{\beta}_{j(k+m)} \left(x_{ji} - K_{jk} \right)_{+}^{m} \right)$$

Titik knot merupakan titik dimana pola dari data berubah. Untuk mendapatkan titik knot optimal maka menggunakan metode GCV. Untuk memilih nilai knot yang paling optimal digunakan nilai GCV yang paling minimum. Titik knot yang digunakan dalam penelitian ini adalah satu knot, dua knot, tiga knot dan kombinasi knot.

4.4.1 Pemilihan Titik Knot Optimal Model Spline dengan Satu Knot

Pemilihan titik knot optimal diawali dengan satu titik knot. Dengan menggunakan satu titik knot pada variabel-variabel yang mempengaruhi kemiskinan diharapkan dapat menemukan GCV yang paling minimum. GCV minimum tersebut diharapkan nantinya dapat menghasilkan model Spline yang terbaik. Berikut ini adalah model regresi nonparametrik Spline *Truncated* dengan satu titik knot pada kasus kemiskinan.

$$\begin{aligned} \hat{y} = {} & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 \left(x_1 - K_1 \right)_+ + \hat{\beta}_3 x_2 + \hat{\beta}_4 \left(x_2 - K_2 \right)_+ + \\ & \hat{\beta}_5 x_3 + \hat{\beta}_6 \left(x_3 - K_3 \right)_+ + \hat{\beta}_7 x_4 + \hat{\beta}_8 \left(x_4 - K_4 \right)_+ + \hat{\beta}_9 x_5 + \hat{\beta}_{10} \left(x_5 - K_5 \right)_+ . \end{aligned}$$

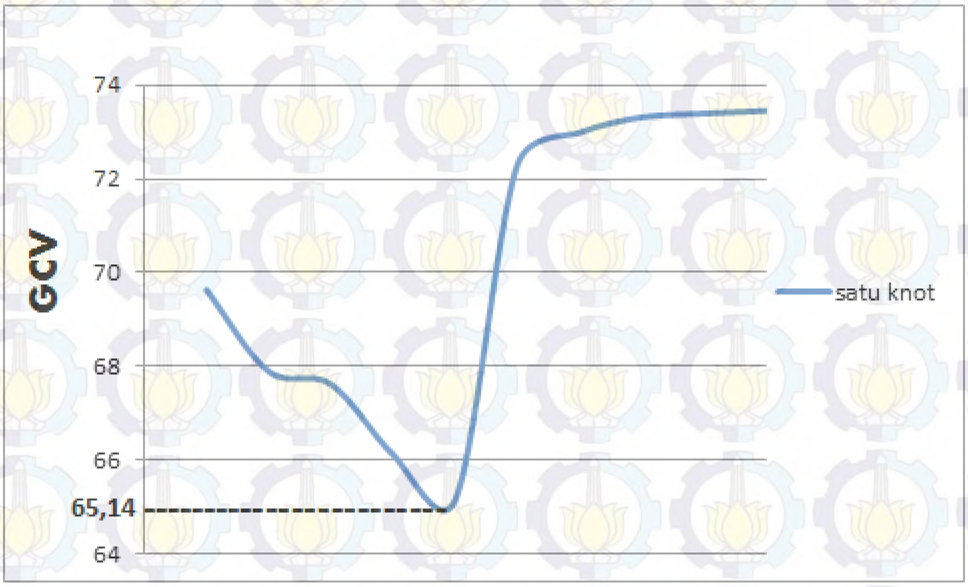
Model regresi nonparametrik Spline *Truncated* yang terbaik diperoleh dari titik-titik knot yang optimal. Untuk mendapatkan titik knot yang optimal, digunakan GCV paling kecil. Berikut adalah hasil perhitungan GCV untuk regresi nonparametrik Spline satu knot.

Tabel 4.1 Pemilihan Titik Knot Optimal dengan Satu Titik Knot

No	x_1	x_2	x_3	x_4	x_5	GCV
1	91,05	9,99	86,04	89,32	92,10	69,63
2	38,33	3,55	22,30	25,25	44,73	67,89
3	41,26	3,91	25,84	28,81	47,36	67,62
4	39,80	3,73	24,07	27,03	46,05	66,12
5	**85,20**	**9,27**	**78,96**	**82,20**	**86,84**	**65,14**
6	89,59	9,81	84,27	87,54	90,79	72,29
7	86,66	9,45	80,73	83,98	88,16	72,98
8	42,72	4,09	27,61	30,59	48,68	73,31
9	88,12	9,63	82,50	85,76	89,47	73,39
10	36,87	3,37	20,53	23,47	43,42	73,45

Berdasarkan Tabel 4.1 dan Gambar 4.6 tersebut terlihat bahwa nilai GCV paling kecil adalah 65,14 dengan titik knot optimal adalah sebagai berikut:

K_1= 85,20, K_2= 9,27, K_3= 78,96, K_4=82,20 dan K_5= 86,84.



Gambar 4.6 GCV untuk Satu Knot

Hasil ini akan dibandingkan dengan menggunakan dua knot, tiga knot, dan kombinasi antara satu, dua, dan tiga knot. Perbandingan ini dilakukan untuk mendapatkan hasil model Spline terbaik.

4.4.2 Pemilihan Titik Knot Optimal Model Spline dengan Dua Knot

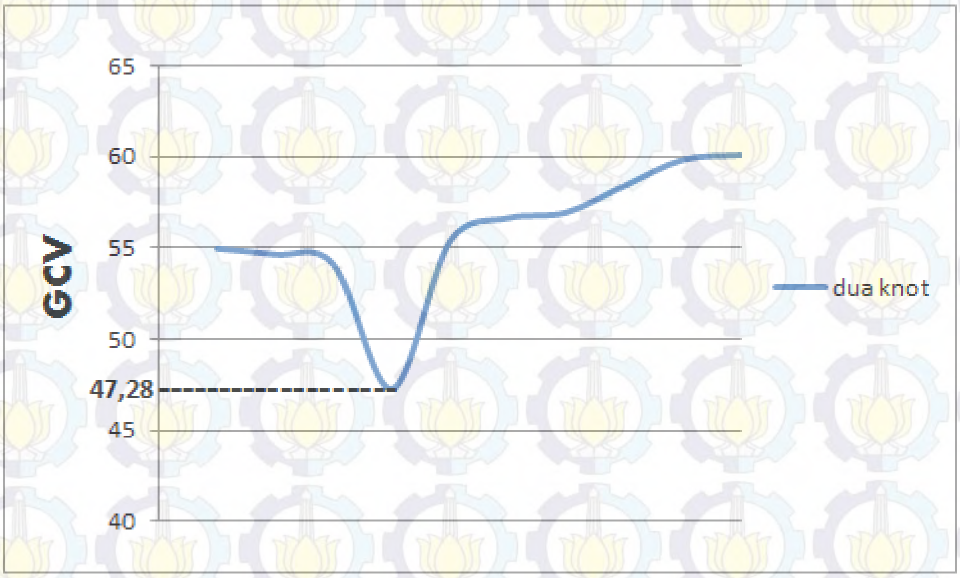
Setelah mendapatkan knot optimal dari satu titik knot maka selanjutnya adalah mencari titik knot optimal dengan dua titik knot. Percobaan dilakukan dengan cara yang serupa dan dipilih GCV yang paling minimum. Berikut adalah model regresi nonparametrik Spline *Truncated* untuk dua titik knot.

$$
\begin{aligned}
\hat{y} = {} & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+ + \hat{\beta}_3 (x_1 - K_2)_+ + \hat{\beta}_4 x_2 + \hat{\beta}_5 (x_2 - K_3)_+ + \hat{\beta}_6 (x_2 - K_4)_+ + \\
& \hat{\beta}_7 x_3 + \hat{\beta}_8 (x_3 - K_5)_+ + \hat{\beta}_9 (x_3 - K_6)_+ + \hat{\beta}_{10} x_4 + \hat{\beta}_{11} (x_4 - K_7)_+ + \hat{\beta}_{12} (x_4 - K_8)_+ + \\
& \hat{\beta}_{13} x_5 + \hat{\beta}_{13} (x_5 - K_9)_+ + \hat{\beta}_{14} (x_5 - K_{10})_+ .
\end{aligned}
$$

Dengan menggunakan persamaan di atas maka dibutuhkan dua titik knot yang optimal untuk masing-masing variabel prediktor. Untuk memperolah knot yang optimal maka dipilih nilai GCV yang paling minimum. Nilai GCV dari dua titik knot diberikan pada tabel 4.2 berikut.

Tabel 4.2 Pemilihan Titik Knot Optimal dengan Dua Titik Knot

No	x_1	x_2	x_3	x_4	x_5	GCV
1	70,55	7,48	61,25	64,40	73,68	54,96
	73,48	7,84	64,79	67,96	76,31	
2	82,27	8,91	75,42	78,64	84,21	54,63
	85,20	9,27	78,96	82,20	86,84	
3	70,55	7,48	61,25	64,40	73,68	54,10
	72,01	7,66	63,02	66,18	75,00	
4	**80,80**	**8,74**	**73,64**	**76,86**	**82,89**	**47,28**
	85,20	**9,27**	**78,96**	**82,20**	**86,84**	
5	79,34	8,56	71,87	75,08	81,58	55,37
	85,20	9,27	78,96	82,20	86,84	
6	72,01	7,66	63,02	66,18	75,00	56,63
	73,48	7,84	64,79	67,96	76,31	
7	35,40	3,19	18,76	21,69	42,10	56,93
	38,33	3,55	22,30	25,25	44,73	
8	77,87	8,38	70,10	73,30	80,26	58,40
	85,20	9,27	78,96	82,20	86,84	
9	76,41	8,20	68,33	71,52	78,95	59,81
	85,20	9,27	78,96	82,20	86,84	
10	48,58	4,80	34,70	37,71	53,94	60,08
	50,05	4,98	36,47	39,49	55,26	



Gambar 4.7 GCV untuk Dua Knot

Berdasarkan Tabel 4.2 dan Gambar 4.7 terlihat bahwa nilai GCV minimum adalah 47,28, dengan titik knot optimal adalah:

(K_1=80,80; K_2=85,20), (K_3=8,74; K_4=9,27),

(K_5=73,64; K_6=78,96), (K_7=76,86; K_8=82,20),

(K_9= 82,89;K_{10}=86,84).

Nilai GCV dengan dua titik knot lebih kecil dari satu titik knot. Sehingga dalam hal ini model nonparametrik Spline *Truncated* dengan dua titik knot lebih baik daripada dengan menggunakan satu titik knot.

4.4.3 Pemilihan Titik Knot Optimal Model Spline dengan Tiga Knot

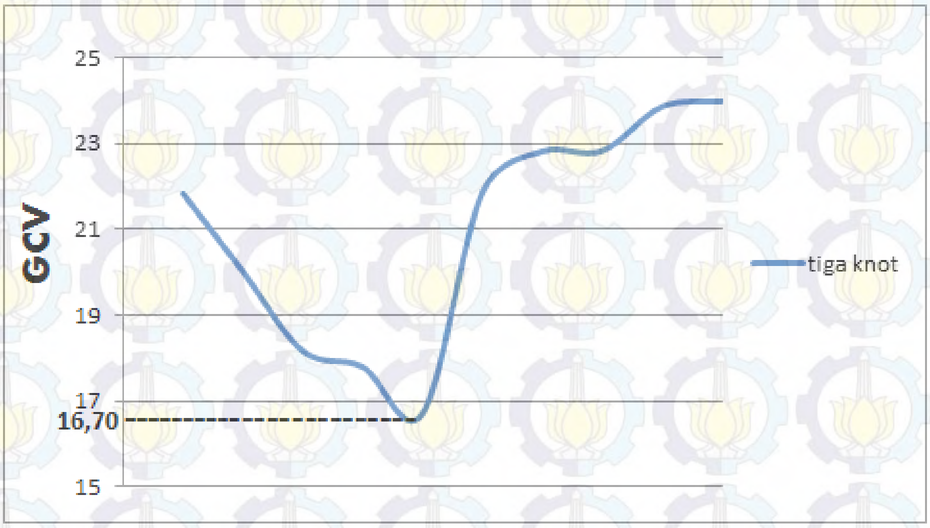
Setelah melakukan pemodelan dengan satu dan dua knot, pada kali ini akan mencoba mencari titik knot optimal dengan menggunakan tiga titik knot. Berikut adalah model regresi nonparametrik Spline *Truncated* dengan tiga titik knot.

$$
\begin{aligned}
\hat{y} = {} & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+ + \hat{\beta}_3 (x_1 - K_2)_+ + \hat{\beta}_4 (x_1 - K_3)_+ + \\
& \hat{\beta}_5 x_2 + \hat{\beta}_6 (x_2 - K_4)_+ + \hat{\beta}_7 (x_2 - K_5)_+ + \hat{\beta}_8 (x_2 - K_6)_+ + \\
& \hat{\beta}_9 x_3 + \hat{\beta}_{10} (x_3 - K_7)_+ + \hat{\beta}_{11} (x_3 - K_8)_+ + \hat{\beta}_{12} (x_4 - K_9)_+ + \\
& \hat{\beta}_{13} x_4 + \hat{\beta}_{14} (x_4 - K_{10})_+ + \hat{\beta}_{15} (x_4 - K_{11})_+ + \hat{\beta}_{16} (x_4 - K_{12})_+ + \\
& \hat{\beta}_{17} x_5 + \hat{\beta}_{18} (x_5 - K_{13})_+ + \hat{\beta}_{19} (x_5 - K_{14})_+ + \hat{\beta}_{20} (x_5 - K_{15})_+ .
\end{aligned}
$$

Untuk membentuk model seperti di atas maka dibutuhkan titik-titik knot optimal. Titik knot optimal tersebut didapatkan dari nilai GCV yang paling minimum. Berikut adalah pemilihan titik knot optimal dengan tiga titik knot.

Tabel 4.3 Pemilihan Titik Knot Optimal dengan Tiga Titik Knot

No	x_1	x_2	x_3	x_4	x_5	GCV
1	61,76	6,41	50,63	53,73	65,79	21,83
	76,41	8,20	68,33	71,52	78,95	
	98,38	10,88	94,89	98,22	98,68	
2	58,83	6,05	47,09	50,17	63,15	20,01
	60,30	6,23	48,86	51,95	64,47	
	89,59	9,81	84,27	87,54	90,79	
3	41,26	3,91	25,84	28,81	47,36	18,16
	63,23	6,59	52,40	55,51	67,10	
	67,62	7,13	57,71	60,84	71,05	
4	61,76	6,41	50,63	53,73	65,79	17,79
	74,94	8,02	66,56	69,74	77,63	
	98,38	10,88	94,89	98,22	98,68	
5	**61,76**	**6,41**	**50,63**	**53,73**	**65,79**	**16,70**
	73,48	**7,84**	**64,79**	**67,96**	**76,31**	
	98,38	**10,88**	**94,89**	**98,22**	**98,68**	
6	63,23	6,59	52,40	55,51	67,10	21,86
	72,01	7,66	63,02	66,18	75,00	
	98,38	10,88	94,89	98,22	98,68	
7	39,80	3,73	24,07	27,03	46,05	22,81
	63,23	6,59	52,40	55,51	67,10	
	67,62	7,13	57,71	60,84	71,05	
8	63,23	6,59	52,40	55,51	67,10	22,82
	73,48	7,84	64,79	67,96	76,31	
	98,38	10,88	94,89	98,22	98,68	
9	61,76	6,41	50,63	53,73	65,79	23,85
	77,87	8,38	70,10	73,30	80,26	
	98,38	10,88	94,89	98,22	98,68	
10	61,76	6,41	50,63	53,73	65,79	23,98
	72,01	7,66	63,02	66,18	75,00	
	98,38	10,88	94,89	98,22	98,68	



Gambar 4.8 GCV untuk Tiga Knot

Berdasarkan Tabel 4.3 dan terlihat pada Gambar 4.8 diketahui bahwa nilai GCV minimum adalah 16,70. Jika dibandingkan dengan percobaan sebelumnya dengan menggunakan satu titik knot dan dua titik knot, model ini memiliki nilai GCV yang lebih kecil. Sehingga dapat dikatakan bahwa model regresi nonprametrik Spline *Truncated* ini lebih baik. Titik-titik knot optimal untuk model ini adalah.

(K_1=61,76; K_2=73,48; K_3=98,38), (K_4=6,41; K_5=7,84; K_6=10,88),

(K_7=50,63; K_8=64,79; K_9=94,89), (K_{10}=53,73; K_{11}=67,96; K_{12}=98,22),

(K_{13}=65,79; K_{14}=76,31; K_{15}=98,68).

4.4.4 Pemilihan Titik Knot Optimal dengan Kombinasi Knot

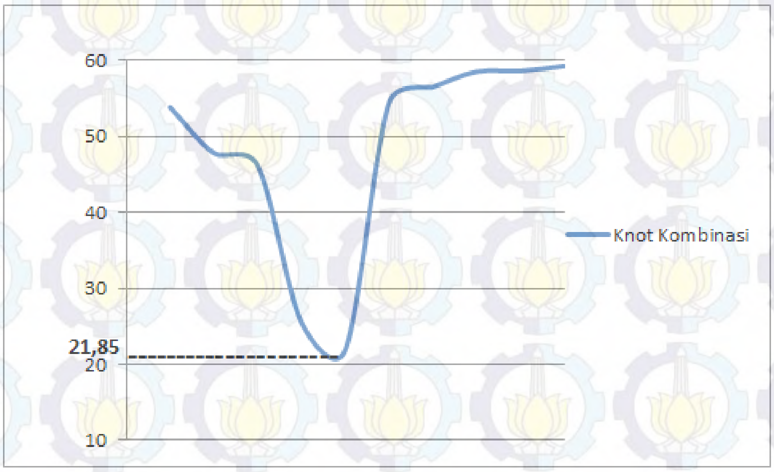
Setelah melakukan pemilihan titik knot optimal dengan satu titik knot, dua titik knot, dan tiga titik knot, selanjutnya melakukan pemilihan titik knot optimal dengan kombinasi knot. Kombinasi knot adalah kombinasi antara satu titik knot, dua titik knot, dan tiga titik knot. Kombinasi ini digunakan untuk memilih titik knot optimal. Dalam pemilihan titik knot optimal pada model regresi nonparametrik Spline *Truncated* dengan kombinasi knot ini dipilih nilai GCV yang paling minimum. Hasil GCV tersebut nantinya akan dibandingkan dengan hasil GCV pada percobaan sebelumnya. Berikut adalah nilai GCV model regresi Spline *Truncated* dengan kombinasi knot.

Tabel 4.4 Pemilihan Titik Knot Optimal dengan Kombinasi Knot

No	Variabel	Variasi Titik Knot	Titik-Titik Knot	GCV
1	x_1	2	K_1=80,80; K_2=85,20	53,84
	x_2	1	K_3 = 9,27	
	x_3	2	K_5 = 73,64; K_6 =78,96	
	x_4	2	K_7 = 76,86; K_8 =82,20	
	x_5	2	K_9 = 21.00; K_{10} =78.90	
2	x_1	2	K_1 =80,80; K_2 =85,20	47,80
	x_2	1	K_3 =9,27	
	x_3	1	K_4 =78,96	
	x_4	2	K_5 =76,86; K_6 =82,20	
	x_5	2	K_7 =21.00; K_8 =78.90	
3	x_1	2	K_1 =80,80; K_2=85,20	46,17
	x_2	2	K_3 =8,74; K_4 =9,27	
	x_3	1	K_5 =78,96	
	x_4	2	K_6 =76,86; K_7 =82,20	
	x_5	2	K_8 =21.00; K_9 =78.90	
4	x_1	3	K_1 =61,76; K_2 =73,48; K_3 =98,38	25,52
	x_2	3	K_4 =6,41; K_5 =7,84; K_6 =10,88	
	x_3	2	K_7 =73,64; K_8 =78,96	
	x_4	3	K_9 =53,73; K_{10} =67,96; K_{11} =98,22	
	x_5	3	K_{12} =65,79; K_{13} =76,31; K_{14} =98,68	
5	x_1	**3**	**K_1 =61,76; K_2 =73,48; K_3 =98,38**	**21,85**
	x_2	**3**	**K_4 =6,41; K_5 =7,84; K_6 =10,88**	
	x_3	**1**	**K_7 =78,96**	
	x_4	**3**	**K_8 =53,73; K_9 =67,96; K_{10} =98,22**	
	x_5	**3**	**K_{11} =65,79; K_{12} =76,31; K_{13} =98,68**	

Tabel 4.4 Pemilihan Titik Knot Optimal dengan Kombinasi Knot (Lanjutan)

No	Variabel	Variasi Titik Knot	Titik-Titik Knot	GCV
6	x_1	2	K_1=80,80; K_2=85,20	54,38
	x_2	3	K_3=6,41; K_4=7,84; K_5=10,88	
	x_3	1	K_6=78,96	
	x_4	2	K_7=76,86; K_8=82,20	
	x_5	2	K_9=21.00; K_{10} =78.90	
7	x_1	2	K_1=80,80; K_2=85,20	56,5
	x_2	3	K_3=6,41; K_4 =7,84; K_5=10,88	
	x_3	2	K_6 =73,64; K_7=78,96	
	x_4	2	K_8 =76,86; K_9=82,20	
	x_5	2	K_{10} =21.00; K_{11}=78.90	
8	x_1	3	K_1=61,76; K_2=73,48; K_3=98,38	58,51
	x_2	2	K_4=8,74; K_5=9,27	
	x_3	3	K_6=50,63; K_7= 64,79; K_8=94,89	
	x_4	3	K_9=53,73; K_{10}=67,96; K_{11}=98,22	
	x_5	3	K_{12}=65,79; K_{13}=76,31; K_{14}=98,68	
9	x_1	2	K_1=80,80; K_2 =85,20	58,62
	x_2	1	K_3=9,27	
	x_3	1	K_4 =78,96	
	x_4	2	K_5=76,86; K_6 =82,20	
	x_5	1	K_7=86,84	
10	x_1	1	K_1=85,20	59,27
	x_2	2	K_2=8,74; K_3 =9,27	
	x_3	1	K_4=78,96	
	x_4	2	K_5=76,86; K_6=82,20	
	x_5	2	K_7=21.00; K_8=78.90	

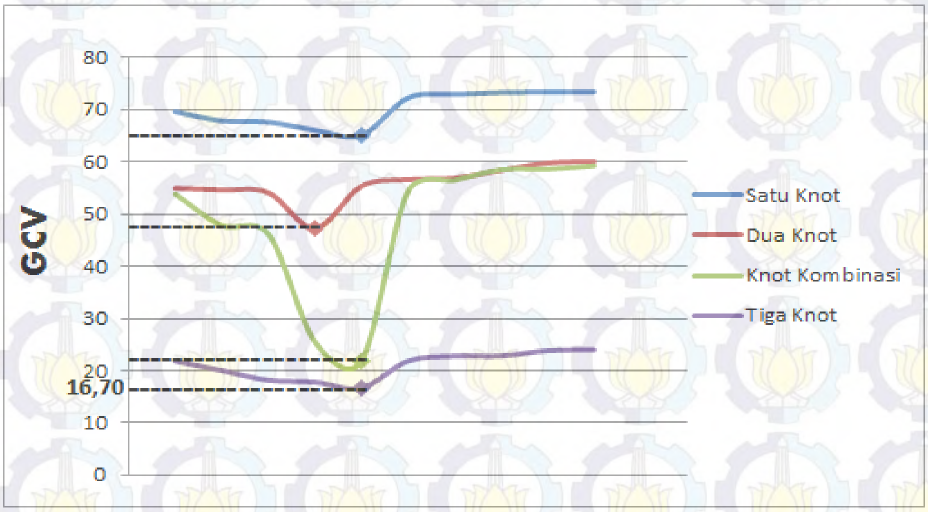


Gambar 4.9 GCV untuk Kombinasi Knot

Pada Tabel 4.4 dan Gambar 4.9 terlihat bahwa nilai GCV minimum dengan kombinasi knot adalah sebesar 21,85 dengan kombinasi 3,3,1,3,3. Titik-titik knot optimal dari kombinasi knot adalah

(K_1=61,76; K_2=73,48; K_3=98,38), (K_4=6,41; K_5=7,84; K_6=10,88), (K_7=78,96), (K_8=53,73; K_9=67,96; K_{10}=98,22), (K_{11}=65,79; K_{12}=76,31; K_{13}=98,68).

Berdasarkan pembahasan yang dilakukan sebelumnya yaitu pemodelan Spline dengan satu, dua, tiga dan kombinasi knot maka dilakukan perbandingan dari model tersebut untuk menentukan model yang terbaik. Ternyata nilai GCV yang paling minimum adalah nilai GCV dengan tiga titik knot. Seperti yang terlihat pada Gambar 4.10 dan Tabel 4.5 berikut ini.



Gambar 4.10 GCV Satu knot, Dua Knot, Tiga Knot dan Kombinasi Knot

Tabel 4.5 Perbandingan Nilai GCV untuk Berbagai Knot

Knot	Nilai GCV
Satu	65,14
Dua	47,28
Tiga	16,70
Kombinasi	21,85

Terlihat pada Gambar 4.10 dan Tabel 4.5 maka nilai yang digunakan untuk analisis selanjutnya adalah nilai dengan tiga titik knot.

4.4.5 Penaksiran Parameter Model Regresi Nonparametrik Spline *Truncated*

Dalam mendapatkan model regresi nonparamterik Spline *Truncated* yang terbaik adalah dengan menggunakan titik knot optimal. Dari hasil pemilihan titik knot optimal yang dilakukan, maka model regresi dengan menggunakan tiga titik knot adalah yang terbaik. Hasil dari estimasi parameter dengan mengunakan tiga titik knot adalah sebagai berikut.

$$\begin{aligned}\hat{y} = {} & 145,11+0,64\,x_1 - 0,83\,(x_1 - 61,76)_+ - 0,39\,(x_1 - 73,48)_+ + 218,58\,(x_1 - 98,38)_+ + \\ & -8,90\,x_2 + 13,97\,(x_2 - 6,41)_+ - 27,72\,(x_2 - 7,84)_+ - 2182,53\,(x_2 - 10,88)_+ + \\ & 0,11\,x_3 + 0,41\,(x_3 - 50,63)_+ - 1,87\,(x_3 - 64,79)_+ + 8,31\,(x_3 - 94,89)_+ + \\ & -5,82\,x_4 + 9,44(x_4 - 53,73)_+ - 1,63(x_4 - 67,96)_+ - 45,01(x_4 - 98,22)_+ + \\ & 3,08\,x_5 - 10\,(x_5 - 65,79)_+ + 5,95\,(x_5 - 76,31)_+ + 70,24\,(x_5 - 98,68)_+ .\end{aligned}$$

Model regresi Spline *Truncated* dengan 3 titik knot ini memiliki R^2 sebesar 98,46%. Hal ini memiliki arti bahwa model ini dapat menjelaskan kemiskinan sebesar 98,46%.

4.4.6 Pengujian Signifikansi Parameter Model Regresi Spline *Truncated*

Pengujian signifikansi parameter ini dilakukan untuk mengetahui apakah parameter yang didapatkan dari hasil pemodelan dengan regresi nonparametrik Spline *Truncated* memiliki pengaruh yang signifikan terhadap presentase kemiskinan. Pengujian signifikansi parameter ini dimulai dengan pengujian parameter secara serentak. Apabila parameter signifikan secara serentak terhadap presentase kemiskinan, maka selanjutnya dilakukan pengujian signifikansi parameter secara individu. Pengujian signifikansi parameter secara individu dilakukan untuk mengetahui signifikansi dari masing-masing parameter terhadap presentase kemiskinan. Langkah pertama yaitu melakukan pengujian signifikansi parameter model secara serentak yang akan dijelaskan pada subbab 4.4.6.1 berikut ini.

4.4.6.1 Pengujian Signifikansi Parameter Model Secara Serentak

Untuk melakukan pengujian signifikansi parameter secara serentak maka menggunakan hipotesis sebagai berikut.

H_0 : $\beta_1 = \beta_2 = ... = \beta_{20} = 0$

H_1 : minimal terdapat satu $\beta_k \neq 0$ dimana $k = 1,2,\ldots,20$

Hasil ANOVA untuk model regresi nonparametrik Spline *Truncated* secara serentak diberikan pada Tabel 4.6 sebagai berikut.

Tabel 4.6 ANOVA Model Regresi Spline Secara Serentak

Sumber	Df	Sum of Square	Mean Square	F_{hitung}	P-value
Regresi	20	2363,076	118,1538	25,64944	0,000035
Error	8	36,85189	4,606486		
Total	28	2399,928			

Berdasarkan hasil ANOVA dapat diketahui bahwa *p-value* adalah sebesar 0,00035. Nilai tersebut kurang dari nilai α (0.05). Sehingga dalam hal ini dapat disimpulkan bahwa H_0 ditolak, maka minimal terdapat satu parameter yang signifikan terhadap variabel respon. Untuk dapat mengetahui parameter mana saja yang berpengaruh terhadap variabel respon maka perlu dilakukan pengujian secara individu. Selanjutnya, akan dilakukan pengujian signifikansi parameter model secara individu yang akan dibahas pada subbab 4.4.6.2 berikut ini.

4.4.6.2 Pengujian Signifikasi Parameter Model Secara Individu

Untuk melakukan pengujian signifikansi parameter secara serentak maka menggunakan hipotesis sebagai berikut.

H_0 : $\beta_k = 0$

H_1 : $\beta_k \neq 0$ dimana k =1,2,…,20

Berikut adalah hasil pengujian signifikansi parameter secara individu.

Tabel 4.7 Pengujian Parameter Model Regresi Secara Individu

Variabel	Parameter	Koefisien	*P-value*	Keputusan
	β_0	145,11	0,000	Signifikan
x_1	β_1	0,64	0,001	Signifikan
	β_2	-0,83	0,098	Tidak Signifikan
	β_3	-0,39	0,427	Tidak Signifikan
	β_4	218,58	0,000	Signifikan
x_2	β_5	– 8,90	0,000	Signifikan
	β_6	13,97	0,001	Signifikan
	β_7	–27,72	0,001	Signifikan
	β_8	–2182,53	0,000	Signifikan
x_3	β_9	0,11	0,506	Tidak Signifikan
	β_{10}	0,41	0,263	Tidak Signifikan
	β_{11}	–1,87	0,018	Signifikan
	β_{12}	8,31	0,117	Tidak Signifikan
x_4	β_{13}	– 5,82	0,000	Signifikan
	β_{14}	9,44	0,000	Signifikan
	β_{15}	–1,63	0,084	Tidak Signifikan
	β_{16}	– 45,01	0,039	Signifikan
x_5	β_{17}	3,08	0,000	Signifikan
	β_{18}	–10,00	0,000	Signifikan
	β_{19}	5,95	0,000	Signifikan
	β_{20}	70,24	0,031	Signifikan

Berdasarkan Tabel 4.7 terlihat bahwa terdapat beberapa parameter yang tidak signifikan. Dari 21 parameter pada model regresi nonparametrik Spline *Truncated*, terdapat 6 parameter yang tidak signifikan. Parameter tersebut tidak signifikan pada taraf signifikansi 0.05, karena *p-value* lebih dari α. Meskipun terdapat 6 parameter yang tidak signifikan, namun secara keseluruhan kelima variabel yang diduga berpengaruh terhadap presentase kemiskinan. Langkah selanjutnya yaitu melakukan pengujian asumsi residual yang akan dibahas pada subbab 4.5.

4.5 Pengujian Asumsi Residual

Pada analisis regresi nonparametrik Spline *Truncated*, residual yang terbentuk harus memenuhi asumsi IIDN. Pengujian asumsi dilakukan untuk mengidentifikasi apakah residual yang terbentuk identik, independen, dan berdistribusi normal. Berikut adalah hasil pengujian asumsi residual.

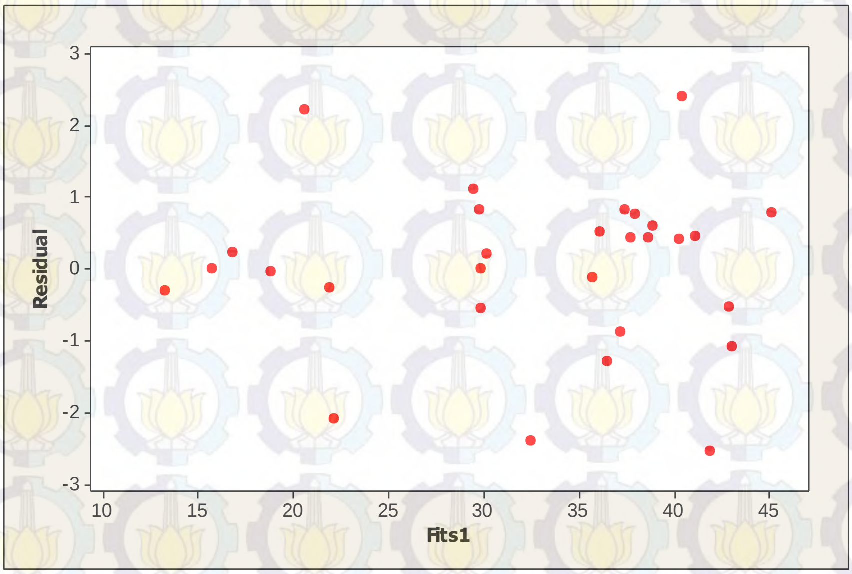
4.5.1 Uji Identik

Untuk melakukan pengujian asumsi residual identik maka menggunakan hipotesis sebagai berikut:

$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_n^2 = \sigma^2$.

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2$, $i = 1, 2, \ldots, n$.

Hipotesis di atas digunakan untuk mengetahui apakah variansi dari residual tersebut identik atau tidak. Untuk mengetahui apakah residual identik atau tidak, dapat dilakukan dengan uji secara visual dan uji *Glejser*. Uji secara visual dilakukan dengan cara membuat plot antara nilai dugaan respon dengan residual. Apabila pada plot tidak terdapat pola maka itu artinya tidak terjadi heterokedastisitas. Berikut adalah hasil plot antara nilai dugaan respon dan residual.



Gambar 4.11 *Scatterplot* antara Fits dan Residual

Berdasarkan plot pada Gambar 4.11 terlihat pencaranya menyebar ke segala arah dan tidak membentuk adanya suatu pola, sehingga secara visual asumsi identik telah terpenuhi. Selain itu untuk mengetahui terjadi heteroskedastisitas dapat dilakukan dengan uji *Glejser*. Uji *Glejser* dilakukan dengan cara meregresikan nilai mutlak dari residual dengan variabel prediktor yang signifikan terhadap model. Berikut adalah hasil dari uji *Glejser*.

Tabel 4.8 ANOVA dari Uji *Glejser*

Sumber	**Df**	***Sum of Square***	***Mean Square***	F_{hitung}	***P-value***
Regresi	20	10,68937	0,5344683	0,7323337	0,7299152
Error	8	5,838522	0,7298152		
Total	28	16,52789			

Dari ANOVA pada Tabel 4.8 dapat diketahui bahwa *p-value* dari uji *Glejser* adalah sebesar 0,7299152, yaitu lebih besar dari nilai α(0,05). Sehingga dapat diputuskan bahwa H_0 gagal ditolak. Jadi dapat diartikan bahwa tidak terjadi heteroskedastisitas. Hal ini menunjukkan residual telah memenuhi asumsi identik. Berikut akan dibahas mengenai uji independen pada subbab 4.5.2.

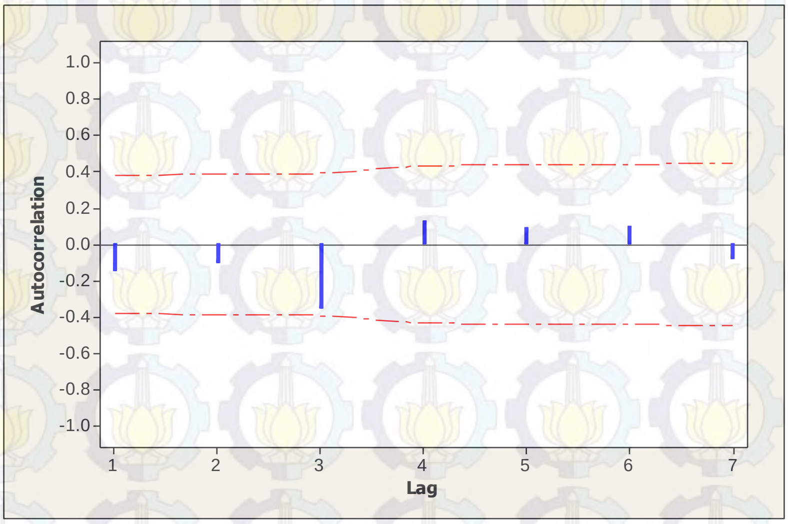
4.5.2 Uji Independen

Hipotesis yang digunakan dalam pengujian asumsi residual independen adalah sebagai berikut.

$H_0 : \rho = 0$ (residual independen)

$H_1 : \rho \neq 0$ (residual tidak independen)

Untuk melakukan apakah residual independen atau tidak dapat dilakukan dengan cara melihat plot ACF dari residual. Apabila terdapat minimal satu autokorelasi pada lag yang keluar dari batas signifikansi maka H_0 ditolak, yaitu residual tidak independen. Berikut adalah hasil plot ACF dari residual.



Gambar 4.12 ACF dari Residual

Berdasarkan plot ACF dari residual pada Gambar 4.12 terlihat bahwa autokorelasi pada semua lag berada di dalam batas atau bisa dikatakan tidak ada autokorelasi yang keluar dari batas. Sehingga dapat disimpulkan bahwa H_0 tidak ditolak, maka residual telah memenuhi asumsi independen. Pengujian asumsi residual yang terakhir adalah menguji distribusi normal. Pengujian tersebut akan dibahas pada subbab 4.5.3 berikut ini.

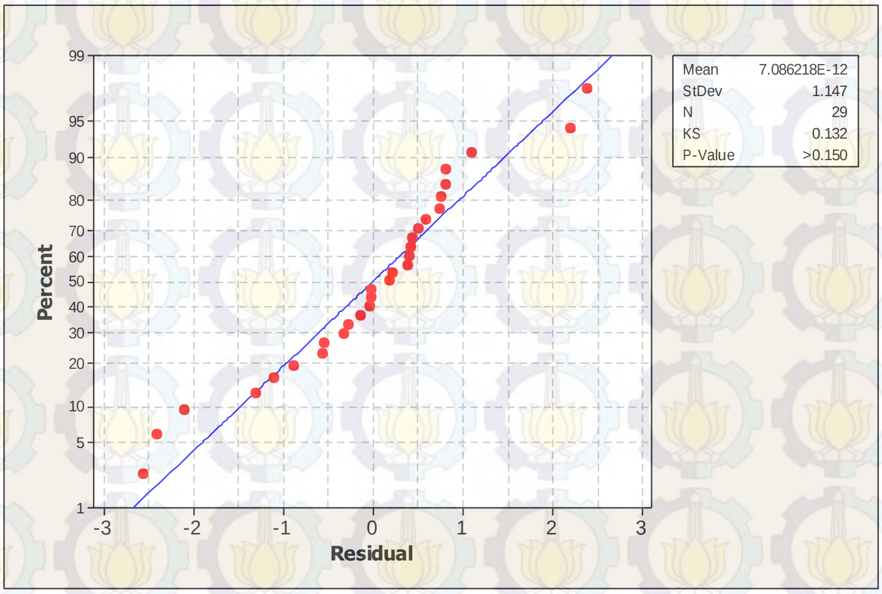
4.5.3 Uji Distribusi Normal

Hipotesis yang digunakan dalam pengujian asumsi residual berdistribusi normal adalah sebagai berikut.

$H_0 : F_0(x) = F(x)$ (residual berdistribusi normal)

$H_1 : F_0(x) \neq F(x)$ (residual tidak berdistribusi normal)

Pengujian distribusi normal dapat dilakukan dengan Uji *Kolmogorov-Smirnov*. Berikut adalah hasil pengujian dengan *Kolmogorov-Smornov*.



Gambar 4.13 Hasil Uji *Kolmogorov-Smirnov*

Berdasarkan Gambar 4.13 terlihat bahwa *p-value* dari Uji *Kolmogorov-Smirnov* menunjukkan nilai >0,150, nilai ini lebih dari nilai α (0,05). Maka dapat diputuskan bahwa gagal tolak H_0. Dapat disimpulkan bahwa residual telah berditribusi normal.

4.6 Interpretasi Model Regresi Nonparametrik Spline *Truncated*

Dari hasil analisis, model regresi nonparametrik Spline *Truncated* terbaik adalah dengan tiga titik knot. Berikut adalah model terbaik yang telah diperoleh.

$$\begin{aligned}\hat{y} = {} & 145{,}11+0{,}64\,x_1 - 0{,}83\,(x_1-61{,}76)_+ - 0{,}39\,(x_1-73{,}48)_+ + 218{,}58\,(x_1-98{,}38)_+ + \\ & -8{,}90\,x_2 + 13{,}97\,(x_2-6{,}41)_+ - 27{,}72\,(x_2-7{,}84)_+ - 2182{,}53\,(x_2-10{,}88)_+ + \\ & 0{,}11\,x_3 + 0{,}41\,(x_3-50{,}63)_+ - 1{,}87\,(x_3-64{,}79)_+ + 8{,}31\,(x_3-94{,}89)_+ + \\ & -5{,}82\,x_4 + 9{,}44(x_4-53{,}73)_+ - 1{,}63(x_4-67{,}96)_+ - 45{,}01(x_4-98{,}22)_+ + \\ & 3{,}08\,x_5 - 10\,(x_5-65{,}79)_+ + 5{,}95\,(x_5-76{,}31)_+ + 70{,}24\,(x_5 - 98{,}68)_+ .\end{aligned}$$

model terbaik tersebut dapat diinterpretasikan sebagai berikut.

1. Apabila variabel x_2, x_3, x_4, dan x_5, dianggap konstan, maka besar pengaruh persentase angka melek huruf (x_1) terhadap presentase kemiskinan (y) adalah

$$\hat{y} = 0{,}64\,x_1 - 0{,}83\,(x_1-61{,}76)_+ - 0{,}39\,(x_1-73{,}48)_+ + 218{,}58\,(x_1-98{,}38)_+$$

$$
= \begin{cases} 0{,}64\, x_1 ; & x_1 < 61{,}76 \\ -0{,}19\, x_1 + 51{,}26 ; & 61{,}76 \le x_1 < 73{,}48 \\ -0{,}58\, x_1 + 79{,}92 ; & 73{,}48 \le x_1 < 98{,}38 \\ 218\, x_1 - 21423{,}98 ; & x_1 \ge 98{,}38 \end{cases}
$$

Dari model tersebut dapat diinterpretasikan yaitu apabila persentase angka melek huruf antara 61,76% hingga 73,48% dan apabila persentase angka melek huruf ini naik 1%, maka presentase kemiskinan akan cenderung turun sebesar 0.0019. Wilayah yang termasuk dalam interval ini adalah kabupaten Paniai dan Membramo Raya. Selanjutnya apabila persentase angka melek huruf berkisar antara 73,48% hingga 98,38% dan apabila persentase angka melek huruf naik 1%, maka Presentase penduduk miskin akan cenderung turun sebesar 0,0058. Wilayah yang masuk dalam interval ini adalah kabupaten Waropen, Nabire, Puncak Jaya, Sarmi, Mimika, Merauke, Yapon Waropen, Keerom, Supiori dan Jayapura. Apabila presentase angka melek huruf lebih dari 98,38 dan apabila presentase angka melek huruf naik 1%, maka presentase kemiskinan akan naik sebesar 2,18. Wilayah yang termasuk dalam interval ini adalah kabupaten Biak Numfor dan Kota Jayapura.

2. Apabila variabel x_1, x_3, x_4, dan x_5 dianggap konstan, maka besar pengaruh rata-rata lama sekolah (x_2) presentase kemiskinan (y) adalah

$$
\hat{y} = -8{,}90\, x_2 + 13{,}97\, (x_2 - 6{,}41)_+ - 27{,}72\, (x_2 - 7{,}84)_+ - 2182{,}53\, (x_2 - 10{,}88)_+
$$

$$
= \begin{cases} -8{,}90\, x_2 ; & x_2 < 6{,}41 \\ 5{,}07\, x_2 - 89{,}55 ; & 6{,}41 \le x_2 < 7{,}84 \\ 22{,}65\, x_2 + 127{,}78 ; & 7{,}84 \le x_2 < 10{,}88 \\ -2205{,}18\, x_2 + 23873{,}70 ; & x_2 \ge 10{,}88 \end{cases}
$$

Dari model tersebut dapat diinterpretasikan yaitu apabila rata-rata lama sekolah kurang dari 6,41 dan apabila rata-rata lama sekolah naik 1%, maka presentase kemiskinan akan cenderung turun sebesar 0,08. Wilayah yang termasuk dalam interval ini adalah Intan Jaya, Pegunungan Bintang, Nduga, Yalimo , Puncak, Yahukimo, Mamberamo Tengah , Deiyai, Tolikara, Boven Digoel, Lanny Jaya, Dogiyai, Mappi, Asmat, Membramo Raya, Jayawijaya, Puncak Jaya, Paniai. Rata-rata lama sekolah antara 6,41 hingga 7,84, apabila rata-rata lama sekolah naik 1%, maka presentase kemiskinan akan cenderung naik sebesar 0,05. Wilayah yang termasuk dalam interval ini adalah kabupaten Waropen, Yapen Waropen, Mimika, Sarmi, Nabire dan Keerom. Selanjutnya apabila rata-rata lama sekolah antara 7,84 hingga 10,88 dan apabila rata-rata lama sekolah naik 1%, maka presentase kemiskinan naik sebesar 0,22. Wilayah yang masuk dalam interval ini adalah kabupaten Supiori, Merauke, Jayapura dan Biak Numfor. Apabila rata-rata lama sekolah lebih dari 10,88 dan apabila rata-rata lama sekolah naik 1%, maka presentase kemiskinan akan berkurang sebesar 22,05. Wilayah yang termasuk dalam interval ini adalah Kota Jayapura.

3. Apabila variabel x_1, x_2, x_4, dan x_5, dianggap konstan, maka besar pengaruh persentase pendidikan kurang dari SD (x_3) terhadap presentase kemiskinan (y) adalah

$$\hat{y} = 0,11\, x_3 + 0,41\,(x_3 - 50,63)_+ - 1,87\,(x_3 - 64,79)_+ + 8,31\,(x_3 - 94,89)_+$$

$$= \begin{cases} 0,11\, x_3 \;; & x_3 < 50,63 \\ 0,52\, x_3 \text{-}20,76 \;; & 50,63 \le x_3 < 64,79 \\ \text{-}1,35\, x_3 + 100,40 \;; & 64,79 \le x_3 < 94,89 \\ 6,96\, x_3 \text{-}688,14 \;; & x_3 \ge 94,89 \end{cases}$$

Dari model tersebut dapat diinterpretasikan yaitu apabila persentase pendidikan kurang dari SD sekolah antara 50,63 hingga 64,79, apabila persentase pendidikan kurang dari SD naik 1%, maka presentase kemiskinan akan cenderung naik sebesar 0,0052. Wilayah yang termasuk dalam interval ini adalah kabupaten Mappi, Paniai, Diyai, Keerom, Jayawijaya dan Intan Jaya. Selanjutnya apabila persentase pendidikan kurang dari SD antara 64,79 hingga 94,89 dan apabila persentase pendidikan kurang dari SD naik 1%, maka presentase kemiskinan turun sebesar 0,01. Wilayah yang masuk dalam interval ini adalah kabupaten Asmat, Sarmi, Tolikara, Lanny Jaya, Memberamo Tengah, Puncak Jaya, Yalimo. Apabila persentase pendidikan kurang dari SD lebih dari 94,89 dan apabila persentase pendidikan kurang dari SD naik 1%, maka presentase kemiskinan akan naik sebesar 0,06. Wilayah yang termasuk dalam interval ini adalah kabupaten Nduga dan Pegunungan Bintang.

4. Apabila variabel x_1, x_2, x_3, dan x_5, dianggap konstan, maka besar pengaruh presentase penduduk yang bekerja di sektor pertanian (x_4) terhadap presentase kemiskinan (y) adalah

$$\hat{y} = -5,82\, x_4 + 9,44(x_4 - 53,73)_+ - 1,63(x_4 - 67,96)_+ - 45,01(x_4 - 98,22)_+$$

$$= \begin{cases} -5,82x_4\ ; & x_4 < 53,73 \\ 3,62x_4 - 507,21\ ; & 53,73 \le x_4 < 67,96 \\ 1,99x_4 - 396,44\ ; & 67,96 \le x_4 < 98,22 \\ -43,02x_4 + 4024,45\ ; & x_4 \ge 98,22 \end{cases}$$

Dari model tersebut dapat diinterpretasikan yaitu apabila presentase penduduk yang bekerja di sektor pertanian kurang dari 53,73 dan apabila presentase penduduk yang bekerja di sektor pertanian naik 1%, maka presentase kemiskinan akan cenderung turun sebesar 0,05. Wilayah yang termasuk dalam interval ini adalah Kota Jayapura, Jayapura, Supiori, Waropen, Mimika, Biak Namfor dan Yapen Waropen. Apabila presentase penduduk yang bekerja di sektor pertanian antara 53,73 hingga 67,96, apabila presentase penduduk yang bekerja di sektor pertanian naik 1%, maka presentase kemiskinan akan cenderung naik sebesar 0,0362. Wilayah yang termasuk dalam interval ini adalah Boven Digoel, Nabire, Puncak, Merauke dan Puncak Jaya. Selanjutnya apabila presentase penduduk yang bekerja di sektor pertanian antara 67,96 hingga 98,22 dan apabila presentase penduduk yang bekerja di sektor pertanian naik 1%, maka presentase kemiskinan naik sebesar 0,0199. Wilayah yang masuk dalam interval ini adalah kabupaten Intan Jaya, Sarmi, Keerom, Asmat, Mappi, Tolikara, Deiyai, Dogiyai, Paniai, Membramo Raya, Yalimo, Memberamo Tengah dan Yahukimo. Apabila persentase presentase penduduk yang bekerja di sektor pertanian lebih dari 98,22 dan presentase penduduk yang bekerja di sektor pertanian naik 1%, maka presentase kemiskinan akan turun sebesar 0,43. Wilayah yang termasuk dalam interval ini adalah Jayawijaya, Pegunungan Bintang, Lanny Jaya dan Nduga.

5. Apabila variabel x_1, x_2, x_3, dan x_4, dianggap konstan, maka besar pengaruh presentase penduduk yang bekerja di sektor informal (x_5) terhadap presentase kemiskinan (y) adalah

$$\hat{y} = 3,08\,x_5 - 10(x_5 - 65,79)_+ + 5,95\,(x_5 - 76,31)_+ + 70,24\,(x_5 - 98,68)_+$$

$$= \begin{cases} 3,08x_5\ ; & x_5 < 65,79 \\ -6.92x_5 + 657,9\ ; & 65,79 \le x_5 < 76,31 \\ -0,97x_5 + 203,86\ ; & 76,31 \le x_5 < 98,68 \\ 69,27x_5 - 6727,43\ ; & x_5 \ge 98,68 \end{cases}$$

Dari model tersebut dapat diinterpretasikan yaitu apabila presentase penduduk yang bekerja di sektor informal antara 65,79 hingga 76,31, apabila presentase penduduk yang bekerja di sektor informal naik 1%, maka presentase kemiskinan akan cenderung turun sebesar 0,0692. Wilayah yang termasuk dalam interval ini adalah kabupaten Nabire dan Puncak Jaya. Selanjutnya apabila presentase penduduk yang bekerja di sektor informal antara 76,31 hingga 98,68 dan apabila presentase penduduk yang bekerja di sektor informal naik 1%, maka presentase kemiskinan turun sebesar 0,0097. Wilayah yang masuk dalam interval ini adalah kabupaten Intan Jaya, Deiyai, Keerom, Sarmi, Asmat, Mappi, Tolikara, Dogiyai, Paniai, Membramo Raya, Yalimo, Yahukimo, Memberamo Tengah dan Jayawijaya. Apabila persentase presentase penduduk yang bekerja di sektor informal lebih dari 98,68 dan apabila presentase penduduk yang bekerja di sektor informal naik 1%, maka presentase kemiskinan akan naik sebesar 0,69. Wilayah yang termasuk dalam interval ini adalah Pegunungan Bintang, Lanny Jaya dan Nduga.

4.7 Model Regresi Nonparametrik Deret Fourier

Berdasarkan deskripsi data menggunakan *scatterplot*, bahwa hubungan antara presentase penduduk miskin (y) dengan angka melek huruf (x_1), rata-rata lama sekolah (x_2), pendidikan kurang dari SD (x_3), bekerja di sektor pertanian (x_4), dan bekerja di sektor informal (x_5) memiliki pola nonparametrik, sehingga model yang digunakan untuk memodelkan presentase kemiskinan di Provinsi Papua adalah model nonparametrik. Model nonparametrik yang digunakan dinyatakan dalam persamaan berikut:

$$y_i = \mu(x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}) + \varepsilon_i$$
$$= \mu(x_{1i}) + \mu(x_{2i}) + \mu(x_{3i}) + \mu(x_{4i}) + \mu(x_{5i}) + \varepsilon_i, \ \ i = 1,2,...,n$$

Berdasarkan hal uraian tersebut, maka model yang digunakan untuk memodelkan presentase penduduk miskin di Provinsi Papua adalah nonparametrik Deret Fourier. Estimasi model dengan pendekatan Deret Fourier diperoleh sebagai berikut:

$$y_i = \beta_0 + b_1 x_{1i} + \sum_{k=1}^{K} \alpha_{k1} \cos kx_{1i} + b_2 x_{2i} + \sum_{k=1}^{K} \alpha_{k2} \cos kx_{2i} + b_3 x_{3i} + \sum_{k=1}^{K} \alpha_{k3} \cos kx_{3i} +$$
$$+ b_4 x_{4i} + \sum_{k=1}^{K} \alpha_{k4} \cos kx_{4i} + b_5 x_{5i} + \sum_{k=1}^{K} \alpha_{k5} \cos kx_{5i}$$

Dalam regresi nonparametrik Deret Fourier, sangat tergantung pada parameter osilasi (K). Parameter osilasi (K) merupakan jumlah dari osilasi gelombang cosinus pada model. Nilai K yang semakin besar akan mengakibatkan model semakin komplek dan osilasi dari kurva estimasi akan semakin rapat serta mengikuti pola data aktual, sehingga bias semakin kecil dan varian semakin besar. Pada penelitian ini dibatasi hingga nilai parameter osilasi (K) yaitu sebanyak 3. Mendapatkan estimator Deret Fourier yang terbaik dalam regresi nonparametrik perlu dilakukan pemilihan parameter osilasi yang optimal. Pemilihan parameter osilasi optimal dilakukan menggunakan metode GCV yang rumusnya dapat dilihat pada persamaan (2.13). Nilai K optimal yang dipilih adalah nilai K yang menghasilkan nilai GCV minimum. Hasil analisis untuk nilai K diberikan pada Tabel 4.9 berikut ini:

Tabel 4.9 Penentuan Nilai Parameter Osilasi (K) Optimal Berdasarkan GCV

Parameter Osilasi (K)	GCV
1	214,27
2	84,73
3	18,79

4.7.1 Regresi Nonparametrik Deret Fourier untuk K=1

Berdasarkan Tabel 4.9, nilai K=1 ini menghasilkan nilai GCV sebesar 214,27. Regresi Nonparametrik Deret Fourier dengan K=1 menghasilkan nilai R^2 sebesar 69,21%. Artinya, keragaman nilai respon telah mampu dijelaskan oleh variabel prediktor sebesar 69,21%. Dengan demikian, estimasi model regresi nonparametrik Deret Fourier untuk model presentase penduduk miskin di Papua diberikan oleh:

$$
\begin{aligned}
\hat{y}_i &= \hat{\beta}_0 + \hat{b}_1 x_{1i} + \hat{\alpha}_{11}\cos x_{1i} + \hat{b}_2 x_{2i} + \hat{\alpha}_{12}\cos x_{2i} + \hat{b}_3 x_{3i} + \hat{\alpha}_{13}\cos x_{3i} + \\
&\quad \hat{b}_4 x_{4i} + \hat{\alpha}_{14}\cos x_{4i} + \hat{b}_5 x_{5i} + \hat{\alpha}_{15}\cos x_{5i} \\
&= 21,88 + 0,01x_{1i} - 1,76\cos x_{1i} - 0,76x_{2i} - 0,59\cos x_{2i} + 0,24x_{3i} + 0,99\cos x_{3i} + \\
&\quad -0,37x_{4i} - 9,30\cos x_{4i} + 0,35x_{5i} + 4,53\cos x_{5i}.
\end{aligned}
$$

Hasil ini akan dibandingkan dengan menggunakan K=2 dan K=3. Perbandingan ini dilakukan untuk mendapatkan hasil model Deret Fourier terbaik.

4.7.2 Regresi Nonparametrik Deret Fourier untuk K=2

Berdasarkan Tabel 4.9, nilai K=2 ini menghasilkan nilai R^2 dan MSE berturut-turut 72,61% dan 22,67. Artinya, keragaman nilai respon telah mampu dijelaskan oleh variabel prediktor sebesar 72,61%. Dengan demikian, estimasi model regresi nonparametrik Deret Fourier untuk model presentase penduduk miskin di Papua diberikan oleh:

$$
\begin{aligned}
\hat{y}_i &= \hat{\beta}_0 + \hat{b}_1 x_{1i} + \hat{\alpha}_{11}\cos x_{1i} + \hat{\alpha}_{21}\cos 2x_{1i} + \hat{b}_2 x_{2i} + \hat{\alpha}_{12}\cos x_{2i} + \hat{\alpha}_{22}\cos 2x_{2i} + \\
&\quad \hat{b}_3 x_{3i} + \hat{\alpha}_{13}\cos x_{3i} + \hat{\alpha}_{23}\cos 2x_{3i} + \hat{b}_4 x_{4i} + \hat{\alpha}_{14}\cos x_{4i} + \hat{\alpha}_{24}\cos 2x_{4i} + \\
&\quad \hat{b}_5 x_{5i} + \hat{\alpha}_{15}\cos x_{5i} + \hat{\alpha}_{25}\cos 2x_{5i}. \\
&= 16,67 - 0,10x_{1i} - 2,31\cos x_{1i} + 1,39\cos 2x_{1i} + 0,85x_{2i} + 0,24\cos x_{2i} + \\
&\quad 1,39\cos 2x_{2i} + 0,34x_{3i} + 0,59\cos x_{3i} + 1,26\cos 2x_{3i} - 0,38x_{4i} + \\
&\quad -10,02\cos x_{4i} - 2,49\cos 2x_{4i} + 0,31x_{5i} + 3,85\cos x_{5i} + 1,77\cos 2x_{5i}.
\end{aligned}
$$

Nilai GCV dengan K=2 lebih kecil dari K=1. Sehingga dalam hal ini model nonparametrik Deret Fourier dengan K=2 lebih baik daripada dengan menggunakan K=1. Berdasarkan Tabel 4.9 dapat dilihat bahwa nilai K optimal adalah 3. Hal ini dikarenakan bahwa pada nilai K = 3 tersebut dihasilkan nilai GCV minimum.

4.7.3 Regresi Nonparametrik Deret Fourier untuk K=3

Berdasarkan Tabel 4.9, nilai K=3 ini menghasilkan nilai R^2 dan MSE berturut-turut 89,20% dan 8,94. Artinya, keragaman nilai respon telah mampu dijelaskan oleh variabel prediktor sebesar 89,20%. Dengan demikian, estimasi model regresi nonparametrik Deret Fourier untuk model presentase penduduk miskin di Papua diberikan oleh:

$$\begin{aligned}\hat{y}_i = \hat{\beta}_0 &+ \hat{b}_1 x_{1i} + \hat{\alpha}_{11}\cos x_{1i} + \hat{\alpha}_{21}\cos 2x_{1i} + \hat{\alpha}_{31}\cos 3x_{1i} + \\ &\hat{b}_2 x_{2i} + \hat{\alpha}_{12}\cos x_{2i} + \hat{\alpha}_{22}\cos 2x_{2i} + \hat{\alpha}_{32}\cos 3x_{2i} + \\ &\hat{b}_3 x_{3i} + \hat{\alpha}_{13}\cos x_{3i} + \hat{\alpha}_{23}\cos 2x_{3i} + \hat{\alpha}_{33}\cos 3x_{3i} + \\ &\hat{b}_4 x_{4i} + \hat{\alpha}_{14}\cos x_{4i} + \hat{\alpha}_{24}\cos 2x_{4i} + \hat{\alpha}_{34}\cos 3x_{4i} + \\ &\hat{b}_5 x_{5i} + \hat{\alpha}_{15}\cos x_{5i} + \hat{\alpha}_{25}\cos 2x_{5i} + \hat{\alpha}_{35}\cos 3x_{5i}.\end{aligned}$$

Estimasi untuk masing-masing koefisien regresi nonparametrik Deret Fourier untuk model presentase penduduk miskin di Papua diberikan pada Tabel 4.10.

Tabel 4.10 Estimasi Parameter Koefisien Regresi

Parameter	Estimator	Parameter	Estimator	Parameter	Estimator
β_0	16,88	α_{22}	-4,95	α_{14}	-13,58
b_1	-0,27	α_{32}	-0,04	α_{24}	-9,10
α_{11}	-0,02	b_3	0,41	α_{34}	-9,48
α_{21}	0,80	α_{13}	6,35	b_5	0,14
α_{31}	-3,66	α_{23}	-2,99	α_{15}	4,79
b_2	3,69	α_{33}	0,49	α_{25}	10,28
α_{12}	-1,36	b_4	-0,32	α_{35}	4,55

Jadi estimasi model presentase penduduk miskin di Papua dengan pemodelan regresi nonparametrik Deret Fourier diberikan oleh:

$$\begin{aligned}\hat{y}_i = {} & 16,88 - 0,27x_{1i} - 0,02\cos x_{1i} + 0,80\cos 2x_{1i} - 3,66\cos 3x_{1i} + \\ & 3,69x_{2i} - 1,36\cos x_{2i} - 4,95\cos 2x_{2i} - 0,04\cos 3x_{2i} + \\ & 0,41x_{3i} + 6,35\cos x_{3i} - 2,99\cos 2x_{3i} + 0,49\cos 3x_{3i} + \\ & -0,32x_{4i} - 13,58\cos x_{4i} - 9,10\cos 2x_{4i} - 9,48\cos 3x_{4i} + \\ & 0,14x_{5i} + 4,79\cos x_{5i} + 10,28\cos 2x_{5i} + 4,55\cos 3x_{5i}.\end{aligned}$$

Berdasarkan hasil pembahasan terhadap model Spline *Truncated* dan Deret Fourier pada kasus kemiskinan di Provinsi Papua maka dapat disimpulkan bahwa model Spline *Truncated* lebih baik daripada Deret Fourier, hal ini dapat dilihat pada Tabel 4.11. Pada Tabel 4.11 nilai GCV dari Spline *Truncated* adalah 16,70 sedangkan nilai GCV dari Deret Fourier adalah 18,79. Sehingga nilai GCV Spline *Truncated* lebih minimum dibandingkan dengan Deret Fourier. Nilai R^2 pada Spline *Truncated* yaitu sebesar 98,46% sedangkan R^2 nilai pada Deret Fourier sebesar 89,20% dan nilai MSE Spline *Truncated* lebih kecil dibandingkan dengan Deret Fourier.

Tabel 4.11 Perbandingan model Deret Fourier dan Spline *Truncated*

Model	GCV	R^2	MSE
Spline *Truncated*	16,70	98,46%	4,61
Deret Fourier	18,79	89,20%	8,94

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan maka dapat diambil kesimpulan sebagai berikut.

1. Estimator kurva regresi nonparametrik multivariabel Spline *Truncated* diperoleh dari optimasi:

   $$\underset{\underset{\sim}{\beta} \in R^{q(K+2)}}{Min} \left\{ \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} f_j\left(x_{ji}\right) \right)^2 \right\}$$

   Optimasi ini menghasilkan estimator untuk kurva regresi Spline *Truncated*:

   $$\hat{\mu}\left(x_{1i}, x_{2i}, ..., x_{pi}\right) = \sum_{j=1}^{p} \hat{f}_j\left(x_{ji}\right)$$

   $$= \sum_{j=1}^{q} \left(\sum_{v=1}^{m} \hat{\beta}_{vj} x_{ji}^{v} + \sum_{k=1}^{r} \hat{\beta}_{j(k+m)} \left(x_{ji} - K_{jk}\right)_{+}^{m} \right)$$

   $$= \sum_{j=1}^{q} \sum_{v=1}^{m} \hat{\beta}_{vj} x_{ji}^{v} + \sum_{j=1}^{q} \sum_{k=1}^{r} \hat{\beta}_{j(k+m)} \left(x_{ji} - K_{jk}\right)_{+}^{m}.$$

   dimana $\hat{\beta}_{vj}$ dan $\hat{\beta}_{j(k+m)}$ $, v = 1,2,...,m; k = 1,2,...,r,$ dan $j = 1,2,...,p$

   Diperoleh dari:

   $\underset{\sim}{\hat{\beta}} = \left(\underset{\sim}{\hat{\beta}}_1', ..., \underset{\sim}{\hat{\beta}}_p'\right)'$ dan

   $$\underset{\sim}{\hat{\beta}}_1 = \left(\beta_{11}, ..., \beta_{m1}, \beta_{1(1+m)}, ..., \beta_{1(r+m)}\right)', ..., \underset{\sim}{\hat{\beta}}_p = \left(\beta_{1p}, ..., \beta_{mp}, \beta_{p(1+m)}, ..., \beta_{p(r+m)}\right)'.$$

2. Estimator kurva regresi nonparametrik multivariabel Deret Fourier diperoleh dari optimasi:

   $$\underset{\underset{\sim}{\beta} \in R^{q(K+2)}}{Min} \left\{ \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{q} f_j\left(x_{ji}\right) \right)^2 \right\}$$

Optimasi ini menghasilkan estimator Deret Fourier:

$$\hat{\mu}\left(x_{1i},x_{2i},...,x_{qi}\right)=\sum_{j=1}^{q}\hat{f}_j\left(x_{ji}\right)=\sum_{j=1}^{q}\left(\hat{b}_j x_{ji}+\frac{1}{2}\hat{\alpha}_{0j}+\sum_{k=1}^{K}\hat{\alpha}_{kj}\cos kx_{ji}\right)$$

dengan $\hat{b}_j,\hat{\alpha}_{0j},\hat{\alpha}_{kj};j=1,2,...,q;k=1,2...,K,$ diberikan oleh persamaan:

$$\underset{\sim}{\hat{\beta}}(K)=\left[\hat{b}_1\quad \frac{1}{2}\hat{\alpha}_{01}\quad \hat{\alpha}_{11}\quad \cdots \quad \hat{\alpha}_{K1}\quad \vdots \quad \cdots \quad \vdots \quad \hat{b}_q\quad \frac{1}{2}\hat{\alpha}_{0q}\quad \hat{\alpha}_{1q}\quad \cdots \quad \hat{\alpha}_{Kq}\right]'$$

$$=\left(X'(K)X(K)\right)^{-1}X'(K)\underset{\sim}{y}.$$

3. Model regresi nonparametrik Spline *Truncated* terbaik adalah dengan tiga titik knot. Berikut adalah model terbaik yang telah diperoleh.

$$\begin{aligned}\hat{y}\ =\ & 145{,}11+0{,}64\,x_1-0{,}83\,(x_1-61{,}76)_+-0{,}39\,(x_1-73{,}48)_++\ 218{,}58\,(x_1-98{,}38)_++\\ & -8{,}90\,x_2+13{,}97\,(x_2-6{,}41)_+-27{,}72\,(x_2-7{,}84)_+-2182{,}53\,(x_2-10{,}88)_++\\ & 0{,}11\,x_3+0{,}41\,(x_3-50{,}63)_+-1{,}87\,(x_3-64{,}79)_++8{,}31\,(x_3-94{,}89)_++\\ & -5{,}82\,x_4+9{,}44(x_4-53{,}73)_+-1{,}63(x_4-67{,}96)_+-45{,}01(x_4-98{,}22)_++\\ & 3{,}08\,x_5-10\,(x_5-65{,}79)_++5{,}95\,(x_5-76{,}31)_++70{,}24\,(x_5\ \text{-}\ 98{,}68)_+.\end{aligned}$$

4. Model regresi nonparametrik Deret Fourier terbaik adalah dengan K=3. Berikut adalah model yang terbaik berdasarkan data kemiskinan di Provinsi Papua.

$$\begin{aligned}\hat{y}_i=\ & 16{,}88-0{,}27x_{1i}-0{,}02\cos x_{1i}+0{,}80\cos 2x_{1i}-3{,}66\cos 3x_{1i}+\\ & 3{,}69x_{2i}-1{,}36\cos x_{2i}-4{,}95\cos 2x_{2i}-0{,}04\cos 3x_{2i}+\\ & 0{,}41x_{3i}+6{,}35\cos x_{3i}-2{,}99\cos 2x_{3i}+0{,}49\cos 3x_{3i}+\\ & -0{,}32x_{4i}-13{,}58\cos x_{4i}-9{,}10\cos 2x_{4i}-9{,}48\cos 3x_{4i}+\\ & 0{,}14x_{5i}+4{,}79\cos x_{5i}+10{,}28\cos 2x_{5i}+4{,}55\cos 3x_{5i}.\end{aligned}$$

5. Berdasarkan pemodelan yang telah dilakukan menggunakan Spline *Truncated* dan Deret Fourier pada kasus kemiskinan di provinsi Papua maka dapat disimpulkan bahwa model Spline *Truncated* lebih baik daripada Deret Fourier. Nilai GCV dari Spline *Truncated* adalah 16,70 sedangkan nilai GCV dari Deret Fourier adalah 18,79. Sehingga nilai GCV Spline *Truncated* lebih minimum dibandingkan dengan Deret Fourier. Nilai R^2 pada Spline *Truncated* yaitu sebesar 98,46% sedangkan R^2 nilai pada Deret Fourier sebesar 89,20% dan nilai MSE Spline *Truncated* lebih kecil dibandingkan dengan Deret Fourier.

6. Variabel-variabel yaitu angka melek huruf, rata-rata lama sekolah, berpendidikan kurang dari SD, bekerja di sektor pertanian dan bekerja di sektor informal memiliki pengaruh yang signifikan terhadap presentase kemiskinan di Provinsi Papua.

5.2 Saran

Berdasarkan hasil analisis dan kesimpulan yang diperoleh, saran yang dapat diberikan adalah:

1. Model Spline *Truncated* pada penelitian ini hanya terbatas pada tiga titik knot, selanjutnya dapat dikembangkan dengan lebih dari tiga knot.
2. Model Deret Fourier dalam penelitian ini hanya terbatas pada parameter osilasi sebanyak tiga, selanjutnya dapat dikembangkan dengan lebih dari tiga parameter osilasi.
3. Penelitian ini terbatas pada penggunaan regresi Spline *Truncated* linier (orde satu). Untuk penelitian selanjutnya dapat dikembangkan dengan regresi orde dua dan tiga.
4. Data pada penelitian ini terbatas pada data *cross section*, untuk selanjutnya dapat dikembangkan dengan data longitudinal yaitu gabungan antara *cross section* dan *time series*.
5. Bagi pemerintah Provinsi Papua diharapkan agar lebih memperhatikan variabel–variabel yang signifikan pada penelitian ini yang akan memberikan suatu nilai tambah untuk peningkatan derajat kesejahteraan.

(...Halaman Ini Sengaja Dikosongkan...)

DAFTAR PUSTAKA

Asrini, L., J. (2012), “Regresi Semiparametrik Deret Fourier”, *Prosiding Seminar Nasional FMIPA Universitas Negeri Surabaya*, hal. 77-80.

Asrini, L., J, and Budiantara, I.,N. (2014), “Fourier Series Semiparametric Regression Models (Case Study: The Production of Law Land Rice Irrigation in Central Java)”. *ARPN Journal of Engineering and Applied Sciences*, Vol. 9, hal. 1501-1506.

Billier, C., and Fahrmeir, L. (2000), “Bayesian varying-coefficient models using adaptive regression Spline, Statistical Modeling”, http://citeseer.ist.psu.edu/biller00bayesian. html.

Bilodeau, M. (1992), “Fourier Smoother and Additive Models”, *The Canadian Journal of Statistics*, Vol.3, hal.257-269.

Bintariningrum, M., F. (2014). *Pemodelan Regresi Nonparametrik Spline Truncated dan Aplikasinya pada Angka Kelahiran Kasar di Surabaya*, Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Budiantara, I N. (2000), *Estimator Spline dalam Regresi Nonparametrik dan Semiparametrik,* Disertasi, Universitas Gajah Mada. Yogyakarta.

Budiantara, I N. (2007), “Inferensi Statistik untuk Model Spline”*, Jurnal Mat-Stat*, Vol. 7, hal.1-14.

Budiantara, I N. (2009), *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*, Institut Teknologi Sepuluh Nopember Surabaya.

Chernichovsky,D. dan Meesok, O.,A. (1985), *Urban-rural Food and Nutrition Consumption Pattern in Indonesia*, PHN Technical Note 85-5, July 1985, World Bank.

Craven, P. and Wahba, G. (1979), “Smoothing Noisy Data with Spline Functions”*, Numerische Mathematics*, Vol. 31, hal.377-403.

Crainiceanu, C. M. Ruppert, D., and Wand, M.P. (2004), “Bayesian Analysis for Penalized Spline Regression Using Wenbugs”*, Statistical Software,* Vol.14, No.14, hal.6-14.

Damayanti,Y. (2013), *Pemodelan Penduduk Miskin di Jawa Timur Menggunakan Metode Geographically Weighted Regression (GWR)*, Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Draper, N.R., and Smith, H. (1992), *Analisis Regresi Terapan*, PT Gramedia Pustaka Utama Jakarta.

Ekasari, D., F. (2012), *Pemodelan SEM dengan Generalized Structured Component Analysis (GSCA) (Studi Kasus: Penentuan Struktur Model Kemiskinan Kabupaten/Kota di Provinsi Jawa Tengah)*, Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Eddy S.S., dan Agung. (2010), "Faktor-faktor yang Mempengaruhi Kemiskinan secara Makro di Lima Belas Provinsi Tahun 2007", *Jurnal dan Organisasi Manajemen,* Vol.6, No.2, hal.89-100.

Eubank,R.L. (1988), *Spline Smoothing and Nonparametric Regression*, Mercel Dekker, New York.

Eubank, R. L. (1999), *Nonparametric Regression and Spline Smoothing,* Second Edition, New York, Marcal Dekker, Inc.

Eubank, R. L., & Thomas, W. (1993), "Detecting Heterocedasticity in Nonparametric Regression". *Journal of the American Statistical Association*, 387-392.

Fadillah. (2010), *Analisis Regresi Jumlah Penduduk Miskin dengan Faktor – Faktor yang Mempengaruhinya di Jawa Timur*, Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Green, P.,J. and Silverman, B.,W. (1994), *Nonparametric Regression and Generalized Linear Models*, Chapman & Hall, London.

Gujarati, D. N. (2004), *Basic Econometric 4th edition*. New York: The Mc Gra Hill Companies.

Haerdle, W. (1990), *Applied Nonparametric Regression*, Cambrige University Press: New York.

Holmes, C.,C., and Mallick B.,K. (2003), "Generalized Nonlinear Modelling With Multivariate Free-Knot Regression *Spline"*, *Journal of the American Statistical Association*, Vol.98, hal.462.

Howell, J. R. (2007), *Analysis Using Smoothing Splines As Implemented in L M E O In R.* Bringham Young University.

Huang, Z. (2003), "Local Asymptotic for Polynomial Spline Regression", *The Annals of Statistics,* Vol.31, No.5, hal.1600-1635.

Kohn, R., Ansley C. F., Tharm, D. (1991), "The Performance of Cross Validation and Maximum Likelihood Estimators of Spline Smooting Parameters", *Journal of The American Statistical Assosiations*, Vol.86, hal.1042-1050.

Kim, Y.J and Gu, C. (2004), "Smoothing Spline Gaussian Regression: More Sealable Computation Via Efficient Approximation", *Royal Statistical Society.* Series B, Vol. 66, No.2, hal.337-356.

Lee, T. C. M. (2004), "Improved Smoothing Spline Regression by Combining Estimatics of Different Smoothness", *Statistics and Probability Letters,* Vol. 67, hal.133-140.

Li, L. (1986), *Nonlinear Waveled-Based Nonparametric Curve Estimation with Consored Data and Inference on Long Memory Processes,* Midingan State University.

Mubarak, R. (2012). *Analisis Regresi Spline Multivariabel untuk Pemodelan Kematian Penderita Demam Berdarah Dengeu (DBD) di Jawa Timur*. Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Pane, R., Budiantara, I.N., Zain, I., dan Otok, B., W. (2014), "Parametric and Nonparametric Estimators in Fourier Series Semiparametric Regression and Their Characteristics". *Applied Mathematical Sciences*, Vol.102, No.8, hal. 5053-5064.

Pintowati,W dan Otok, B., W. (2012). "Pemodelan Kemiskinan di Provinsi Jawa Timur dengan Pendekatan Multivariate Adaptive", *Jurnal Sains dan Seni ITS*, Vol.1, No.1.

Prahutama, A. (2013), "Model Regresi Nonparametrik dengan Pendekatan Deret Fourier pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur", *Prosiding Seminar Nasional Statistika Undip*, Vol. 10, hal. 69-76.

Prasetyawan, I., F. (2011), *Penentuan Matriks Pembobot yang Optimum pada Pemodelan Geographically Weighted Regression (GWR) (Studi Kasus: Penyusunan Model Kemiskina di Jawa Tengah*, Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Rencher, A, C. (2002), *Methods of Multivariate Analysis.Second Edition,* John Wiley &Sons, Inc.New York.

Rencher, A., dan Schaalje, G.B. (2008), *Linear Models in Statistics 2nd Edition*, John Willey and Sons Inc., New Jersey.

Rusdarti dan Karolina. (2013),"Faktor-faktor yang Mempengaruhi Tingkat Kemiskinan di Provinsi Jawa Tengah", *Jurnal Economia,*Vol.9, No.1.

Samsodin, M. (2012), *Regresi Spline Polynomial Truncated Multirespon untuk Pemodelan Indikator Kemiskinan di Provinsi Jawa Timur*, Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Searle, S.R. (1971), *Linear Models*, John Willey and Sons Inc., New York.

Semiati, R. (2010), *Regresi Nonparametrik Deret Fourier Birespon*, Tesis, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Seruni P.,S., Putu, dan Sutrisna, Ketut. (2014), "Pengaruh PDRB per Kapita, Pendidikan dan Produktivitas Tenaga Kerja terhadap Kemiskinan di Provinsi Bali", *E-Jurnal EP Unud,* Vol.3, No.10, hal. 431-439

Shao, J. (1993), "Linear Model Selection by Cross Validation", *Journal of the American Statistical Association*, Vol. 88, hal 486-494.

Surya, L. (2013), *Faktor – Faktor yang Mempengaruhi Persentase Penduduk Miskin di Jawa Timur Menggunakan Regresi Semiparametrik Spline*, Tugas Akhir, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Tjahjono, E. (2009), *Estimator Deret Fourier Terbobot pada Regresi Nonparametrik*, Tesis, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.

Tripena, A. and Budiantara, I N. (2007), "Fourier Estimator in Nonparametric Regression", *International Conference On Natural Sciences and Applied Natural Scienes*, Ahmad Dahlan University, Yogyakarta.

Tripena, A. (2013), "Estimator Deret Fourier untuk Estimasi Kurva Regresi Nonparametrik Birespon", *Magistra*, Vol. 84, hal. 1-15.

Venter, J.H., and Snyman, J.L.J. (1995), "A Note on the Generalized Cross Validation Criterion in Linear Model Selection", *Biometrika*, Vol.82, hal.215-219.

Wahba, G. (1985), "A Comparizon of GVC and GML for Choosing The Smoothing Parameter in Generalized Spline Smoothing Problem", The *Annal of Statistic*, Vol.13, hal.1378-1402.

Wahba, G. (1990), "Spline Models For Observational Data", *SIAM, CBMS-NSF Regional Conference Series in Applied mathematics*, Philadelphia.

Walpole. (1995), *Pengantar Statistika.* Jakarta: PT. Gramedia Utama.

Wei, W. W. (2006), *Time Series Analisis: Univariate and Multivariate.* USA: Pearson Education Inc.

BIODATA PENULIS

Mirah, nama panggilan akrab dari penulis. Penulis memiliki nama lengkap Ni Putu Ayu Mirah Mariati. Penulis lahir di Gianyar Bali, pada tanggal 17 Maret tahun 1991. Jenjang pendidikan yang telah ditempuh penulis adalah TK Saraswati 3 Denpasar pada tahun 1995-1997. Menempuh pendidikan dasar di SD Saraswati 5 Denpasar pada (1997-2003), pendidikan menengah pertama di SMP Negeri 1 Denpasar (2003-2006),pendidikan menengah atas di SMA Negeri 1 Denpasar (2006-2009).
Setelah lulus SMA penulis melanjutkan jenjang perguruan tinggi S1 di Jurusan Matematika Universitas Udayana pada (2009-2013). Setelah lulus dari jenjang S1, penulis mempunyai kesempatan melanjutkan pendidikan pascasarjana program magister melalui program beasiswa BPPDN calon dosen DIKTI di Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

Segala saran dan kritik yang membangun, sangat penulis harapkan untuk kebaikan ke depannya. Penulis dapat dihubungi di ayumirahmariati@gmail.com atau ayumirahmariati@yahoo.com.

DAFTAR LAMPIRAN

Halaman

(...Halaman Ini Sengaja Dikosongkan...)

Lampiran 1 :Data Kemiskinan di Provinsi Papua dan Faktor-Faktor yang Diduga Berpengaruh

No	Kabupaten/Kota	y	x_1	x_2	x_3	x_4	x_5
1	Merauke	12.95	88.22	9.46	14.41	64.37	64.37
2	Jayawijaya	39.05	52.77	5.31	61.48	98.49	98.49
3	Jayapura	17.08	96.9	9.56	16.48	31.02	37.44
4	Nabire	30.65	83.67	7.29	39.64	59.21	67.01
5	Yapen Waropen	30.35	90.87	6.74	22.32	48.89	63.34
6	Biak Numfor	29.84	98.68	9.64	25.77	48.06	52.82
7	Paniai	38.69	62.95	6.22	58.53	90.85	91.86
8	Puncak Jaya	38.21	86.82	6.12	79.65	67.79	68.23
9	Mimika	20.09	88.2	6.93	33.07	41.68	45.94
10	Boven Digoel	22.79	35.25	3.65	32.82	58.39	57.59
11	Mappi	29.3	33.47	4.36	56.05	84.11	83.73
12	Asmat	30.56	31.15	4.42	65.98	83.06	83.01
13	Yahukimo	41.98	32.77	2.92	77.62	97.84	97.84
14	Pegunungan Bintang	35.63	32.5	2.59	96.66	99.48	99.48
15	Tolikara	36.3	33.45	3.38	69.04	84.22	84.22
16	Sarmi	18.82	87.68	7.00	68.65	81.75	81.75
17	Keerom	21.65	92.39	7.43	60.42	82.07	81.67
18	Waropen	36.63	78.27	6.55	29.89	34.36	35.52
19	Supiori	41.57	96.69	8.10	25.44	33.3	38.53
20	Membramo Raya	35.2	65.36	5.20	76	92.45	93.33
21	Nduga	38.13	30.54	2.81	95.44	100	100
22	Lanny Jaya	42.32	36.92	3.75	75.87	99.53	99.53
23	Mamberamo Tengah	42.83	34.53	2.93	79.6	97.6	98.38
24	Yalimo	39.49	33.52	2.81	80.8	96.99	97.5
25	Puncak	39.38	32.15	2.85	76.25	61.15	61.55
26	Dogiyai	30.08	34.65	4.16	35.76	89.93	90.87
27	Intan Jaya	40.64	28.08	2.30	62.53	76.66	76.66
28	Deiyai	45.92	31.02	2.96	59.19	84.46	81.62
29	Kota Jayapura	15.77	99.84	11.06	9.91	12.79	37.2

Keterangan:
y = Presentase Kemiskinan
x_1 = Angka Melek Huruf
x_2 = Rata-rata Lama Sekolah
x_3 = Pendidikan kurang dari Sekolah Dasar (SD)
x_4 = Bekerja di Sektor Pertanian
x_5 = Bekerja di Sektor Informal

Lampiran 2 :Program untuk Mencari Titik Knot Optimal pada Model Regresi Nonparametrik Spline *Truncated* dengan 1 Knot

```
data=read.table("d://TESIS.txt",header=FALSE)
GCV1=function(para)
{
  data=read.table("d://TESIS.txt",header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  dataA=data[,(para+2):q]
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot1=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
      knot1[j,i]=a[j]
    }
  }
  a1=length(knot1[,1])
  knot1=knot1[2:(a1-1),]
  aa=rep(1,p)
  data1=matrix(ncol=m,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
    for (j in 1:m)
    {
      for (k in 1:p)
      {
        if (data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else
data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
      }
    }
mx=cbind(aa,data2,data1)
    mx=as.matrix(mx)
    C=pinv(t(mx)%*%mx)
    B=C%*%(t(mx)%*%data[,1])
    yhat=mx%*%B
    SSE=0
    SSR=0
    for (r in (1:p))
    {
      sum=(data[r,1]-yhat[r,])^2
      sum1=(yhat[r,]-mean(data[,1]))^2
      SSE=SSE+sum
```

```
SSR=SSR+sum1
    }
    Rsq[i]=(SSR/(SSE+SSR))*100
    MSE=SSE/p
    A=mx%*%C%*%t(mx)
    A1=(F-A)
    A2=(sum(diag(A1))/p)^2
    GCV[i]=MSE/A2
  }
  GCV=as.matrix(GCV)
  Rsq=as.matrix(Rsq)
cat("======================================","\n")
cat("Nilai Knot dengan Spline linear 1 knot","\n")
cat("======================================","\n")
print (knot1)
cat("=======================================","\n")
cat("Rsq dengan Spline linear 1 knot","\n")
cat("=======================================","\n")
print (Rsq)
cat("=======================================","\n")
cat("HASIL GCV dengan Spline linear 1 knot","\n")
cat("=======================================","\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))
cat("======================================","\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot","\n")
cat("======================================","\n")
cat(" GCV =",s1,"\n")
write.csv(GCV,file="d:/output GCV1.csv")
write.csv(Rsq,file="d:/output Rsq1.csv")
write.csv(knot1,file="d:/output knot1.csv")
}
```

Lampiran 3 :Program untuk Mencari Titik Knot Optimal pada Model Regresi Nonparametrik Spline *Truncated* dengan 2 Knot

```
data=read.table("d://TESIS.txt",header=FALSE)
GCV2=function()
{
  data=read.table("d:/TESIS.txt", header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-1
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(data[,(i+1)]),max(data[,(i+1)]),length.out=50)
      knot[j,i]=a[j]
    }
  }
z=(nk*(nk-1)/2)
knot2=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot1=rbind(rep(NA,2))
      for ( j in 1:(nk-1))
      {
            for (k in (j+1):nk)
            {
            xx=cbind(knot[j,i],knot[k,i])
            knot1=rbind(knot1,xx)
            }
      }
knot2=cbind(knot2,knot1)
}
knot2=knot2[2:(z+1),2:(2*m+1)]
  aa=rep(1,p)
  data2=matrix(ncol=(2*m),nrow=p)
  data1=data[,2:q]
  a1=length(knot2[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)
  for (i in 1:a1)
  {
   for (j in 1:(2*m))
      {
      if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
      for (k in 1:p)
      {
      if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else
      data2[k,j]=data1[k,b]-knot2[i,j]
```

Lanjutan. Lampiran 3

```
    }
        }
    mx=cbind(aa,data1,data2)
    mx=as.matrix(mx)
    C=pinv(t(mx)%*%mx)
    B=C%*%(t(mx)%*%data[,1])
    yhat=mx%*%B
    SSE=0
    SSR=0
    for (r in (1:p))
    {
      sum=(data[r,1]-yhat[r,])^2
      sum1=(yhat[r,]-mean(data[,1]))^2
      SSE=SSE+sum
      SSR=SSR+sum1
    }
    Rsq[i]=(SSR/(SSE+SSR))*100
    MSE=SSE/p
    A=mx%*%C%*%t(mx)
    A1=(F-A)
    A2=(sum(diag(A1))/p)^2
    GCV[i]=MSE/A2
    }
GCV=as.matrix(GCV)
  Rsq=as.matrix(Rsq)
  cat("=============================================================","\n")
  cat("Nilai Knot dengan Spline linear 2 knot","\n")
  cat("=============================================================","\n")
  print (knot2)
  cat("=============================================================","\n")
  cat("Rsq dengan Spline linear 2 knot","\n")
  cat("=============================================================","\n")
  print (Rsq)
  cat("=============================================================","\n")
  cat("HASIL GCV dengan Spline linear 2 knot","\n")
  cat("=============================================================","\n")
  print (GCV)
  s1=min(GCV)
  cat("=============================================================","\n")
  cat("HASIL GCV terkecil dengan Spline linear 2 knot","\n")
  cat("=============================================================","\n")
  cat(" GCV =",s1,"\n")
write.csv(GCV,file="d:/output GCV2.csv")
write.csv(Rsq,file="d:/output Rsq2.csv")
write.csv(knot2,file="d:/output knot2.csv")
}
```

Lampiran 4 :Program untuk Mencari Titik Knot Optimal pada Model Regresi Nonparametrik Spline *Truncated* dengan 3 Knot

```
data=read.table("d://TESIS.txt",header=FALSE)
GCV3=function(para)
{
  data=read.table("d://TESIS.txt",header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[,(para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
      knot[j,i]=a[j]
    }
  }
knot=knot[2:(nk-1),]
a2=nrow(knot)
z=(a2*(a2-1)*(a2-2)/6)
knot1=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot2=rbind(rep(NA,3))
      for ( j in 1:(a2-2))
      {
            for (k in (j+1):(a2-1))
            {
                  for (g in (k+1):a2)
                  {
                  xx=cbind(knot[j,i],knot[k,i],knot[g,i])
                  knot2=rbind(knot2,xx)
                  }
            }
      }
knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
  aa=rep(1,p)
  data1=matrix(ncol=(3*m),nrow=p)
  data2=data[,(para+2):q]
  a1=length(knot1[,1])
  GCV=rep(NA,a1)
  Rsq=rep(NA,a1)
for (i in 1:a1)
  {
```

Lanjutan. Lampiran 4

```
for (j in 1:ncol(knot1))
      {
      b=ceiling(j/3)
      for (k in 1:p)
      {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else
      data1[k,j]=data2[k,b]-knot1[i,j]
      }
        }
    mx=cbind(aa,data[,2:q],data1)
    mx=as.matrix(mx)
    C=pinv(t(mx)%*%mx)
    B=C%*%(t(mx)%*%data[,1])
    yhat=mx%*%B
    SSE=0
    SSR=0
    for (r in (1:p))
    {
      sum=(data[r,1]-yhat[r,])^2
      sum1=(yhat[r,]-mean(data[,1]))^2
      SSE=SSE+sum
      SSR=SSR+sum1
    }
    Rsq[i]=(SSR/(SSE+SSR))*100
    MSE=SSE/p
    A=mx%*%C%*%t(mx)
    A1=(F-A)
    A2=(sum(diag(A1))/p)^2
    GCV[i]=MSE/A2
     }
  GCV=as.matrix(GCV)
  Rsq=as.matrix(Rsq)
cat("=======================================","\n")
cat("Nilai Knot dengan Spline linear 3 knot","\n")
cat("=======================================","\n")
print (knot1)
cat("=======================================","\n")
cat("Rsq dengan Spline linear 3 knot","\n")
cat("=======================================","\n")
print (Rsq)
r=max(Rsq)
print (r)
cat("=======================================","\n")
cat("HASIL GCV dengan Spline linear 3 knot","\n")
cat("=======================================","\n")
print (GCV)
s1=min(GCV)
cat("=======================================","\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot","\n")
cat("=======================================","\n")
cat(" GCV =",s1,"\n")
write.csv(GCV,file="d:/output GCV3.csv")
write.csv(Rsq,file="d:/output Rsq3.csv")
write.csv(knot1,file="d:/output knot3.csv")
}
```

Lampiran 5 :Program untuk Mencari Titik Knot Optimal pada Model Regresi Nonparametrik Spline *Truncated* dengan Knot Kombinasi

```
data=read.table("d://TESIS.txt",header=FALSE)
GCVkom=function(para)
{
  data=as.matrix(data)
  p1=length(data[,1])
  q1=length(data[1,])
  v=para+2
  F=matrix(0,nrow=p1,ncol=p1)
  diag(F)=1
x1=read.table("d:/x1.txt")
x2=read.table("d:/x2.txt")
x3=read.table("d:/x3.txt")
x4=read.table("d:/x4.txt")
x5=read.table("d:/x5.txt")
n2=nrow(x1)
a=matrix(nrow=5,ncol=3^5)
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
for (s in 1:3)
{
m=m+1
a[,m]=c(i,j,k,l,s)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^5)
for (i in 1:3^5)
{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,1]==2)
{
gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
```

Lanjutan. Lampiran 5

```
 for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[,(v+1)])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[,(v+1)],data[,(v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[,(v+1)],data[,(v+1)],data[,(v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )
```

Lanjutan. Lampiran 5

```
gen=as.matrix(data[,(v+2)])
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{
gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[,(v+2)],data[,(v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[,(v+2)],data[,(v+2)],data[,(v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[,(v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,4]==2)
{
gab=as.matrix(x4[,2:3] )
gen=as.matrix(cbind(data[,(v+3)],data[,(v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
```

Lanjutan. Lampiran 5

```
}
else
{
gab=as.matrix(x4[,4:6])
gen=as.matrix(cbind(data[,(v+3)],data[,(v+3)],data[,(v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,5]==1)
{
gab=as.matrix(x5[,1] )
gen=as.matrix(data[,(v+4)])
ee=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,5]==2)
{
gab=as.matrix(x5[,2:3] )
gen=as.matrix(cbind(data[,(v+4)],data[,(v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x5[,4:6])
gen=as.matrix(cbind(data[,(v+4)],data[,(v+4)],data[,(v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
ma=as.matrix(cbind(aa,bb,cc,dd,ee))
mx=cbind(rep(1,nrow(data)),data[,2:q1],na.omit(ma))
 mx=as.matrix(mx)
     C=pinv(t(mx)%*%mx)
     B=C%*%(t(mx)%*%data[,1])
     yhat=mx%*%B
SSE=0
```

Lanjutan. Lampiran 5

```
    SSR=0
for (r in 1:nrow(data))
    {
      sum=(data[r,1]-yhat[r,])^2
      sum1=(yhat[r,]-mean(data[,1]))^2
      SSE=SSE+sum
      SSR=SSR+sum1
    }
    Rsq=(SSR/(SSE+SSR))*100
    MSE=SSE/p1
    A=mx%*%C%*%t(mx)
    A1=(F-A)
    A2=(sum(diag(A1))/p1)^2
    GCV[h,i]=MSE/A2
}

if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
if (a[i,5]==1) splines=x5[,1] else
if (a[i,5]==2) splines=x5[,2:3] else
splines=x5[,4:6]
kkk=cbind(sp,spl,splin,spline,splines)
cat("====================","\n")
print(i)
print(kkk)
print(Rsq)
}
write.csv(GCV,file="d:/output GCV kombinasi mirah.csv")
write.csv(Rsq,file="d:/output Rsq kombinasi mirah.csv")
}
```

Lampiran 6 :Program untuk Uji Glejser dengan 3 Knot

```
data=read.table("d://TESIS.txt",header=FALSE)
res=read.table("d://res.txt",header=FALSE)
knot=read.table("d://knot.txt",header=FALSE)

glejser=function(data,knot,res,alpha,para)
{
data=as.matrix(data)
knot=as.matrix(knot)
res=abs(res)
res=as.matrix(res)
rbar=mean(res)
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],da
ta[,m+1],data[,m+2],data[,m+2],data[,m+2],data[,m+3],data[,m+3]
,data[,m+3],
data[,m+4],data[,m+4],data[,m+4])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
      for(j in 1:p)
      {
      if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else
data.knot[j,i]=dataA[j,i]-knot[1,i]
      }
}
mx=cbind(satu,
data[,2],data.knot[,1:3],data[,3],data.knot[,4:6],data[,4],data
.knot[,7:9],data[,5],data.knot[,10:12],data[,6],data.knot[,13:1
5])
mx=as.matrix(mx)
B=(ginv(t(mx)%*%mx))%*%t(mx)%*%res
n1=nrow(B)
yhat=mx%*%B
residual=res-yhat
SSE=sum((res-yhat)^2)
SSR=sum((yhat-rbar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/SST)*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
```

Lanjutan. Lampiran 6

```
{
cat("-------------------------------------","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-------------------------------------","\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan atau terjadi heteroskedastisitas","\n")
cat("","\n")
}
else
{
cat("-------------------------------------","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-------------------------------------","\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan atau tidak terjadi heteroskedastisitas","\n")
cat("","\n")
}
cat("Analysis of Variance","\n")
cat("=======================================","\n")
      cat("Sumber        df       SS        MS        Fhit","\n")
      cat("Regresi      ",(n1-1)," ",SSR," ",MSR,"",Fhit,"\n")
      cat("Error        ",p-n1," ",SSE,"",MSE,"\n")
      cat("Total        ",p-1," ",SST,"\n")
cat("=======================================","\n")
      cat("s=",sqrt(MSE),"    Rsq=",Rsq,"\n")
      cat("pvalue(F)=",pvalue,"\n")
}
```

Lampiran 7 :Program untuk Uji Parameter dengan 3 Knot

```
uji=function(alpha,para)
{
data=read.table("d:/TESIS.txt")
knot=read.table("d:/knot.txt")
data=as.matrix(data)
knot=as.matrix(knot)
ybar=mean(data[,1])
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m],data[,m],data[,m+1],data[,m+1],da
ta[,m+1],data[,m+2],data[,m+2],data[,m+2],data[,m+3],data[,m+3]
,data[,m+3],data[,m+4],data[,m+4],data[,m+4])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
      for(j in 1:p)
      {
      if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else
data.knot[j,i]=dataA[j,i]-knot[1,i]
      }
}
mx=cbind(satu,
data[,2],data.knot[,1:3],data[,3],data.knot[,4:6],data[,4],data
.knot[,7:9],data[,5],data.knot[,10:12],data[,6],data.knot[,13:1
5])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=======================================","\n")
cat("Estimasi Parameter","\n")
cat("=======================================","\n")
print (B)
n1=nrow(B)
yhat=mx%*%B
res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)
SSR=sum((yhat-ybar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/(SSR+SSE))*100
#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("------------------------------------","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("------------------------------------","\n")
```

Lanjutan. Lampiran 7

```
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan","\n")
cat("","\n")
}
else
{
cat("-------------------------------------","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-------------------------------------","\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan","\n")
cat("","\n")
}

#uji t (uji individu)

thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*%mx))))
cat("-------------------------------------","\n")
cat("Kesimpulan hasil uji individu","\n")
cat("-------------------------------------","\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
thit[i]=B[i,1]/SE[i]
pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan
dengan pvalue",pval[i],"\n") else cat("Gagal tolak Ho yakni
prediktor tidak signifikan dengan pvalue",pval[i],"\n")
}
thit=as.matrix(thit)
cat("========================================","\n")
cat("nilai t hitung","\n")
cat("========================================","\n")
print (thit)
cat("Analysis of Variance","\n")
cat("======================================","\n")
      cat("Sumber        df        SS        MS        Fhit","\n")
      cat("Regresi      ",(n1-1)," ",SSR," ",MSR,"",Fhit,"\n")
      cat("Error        ",p-n1," ",SSE,"",MSE,"\n")
      cat("Total        ",p-1," ",SST,"\n")
cat("======================================","\n")
      cat("s=",sqrt(MSE),"    Rsq=",Rsq,"\n")
      cat("pvalue(F)=",pvalue,"\n")
write.csv(res,file="d:/output uji residual.csv")
write.csv(pval,file="d:/output uji pvalue.csv")
write.csv(mx,file="d:/output uji mx.csv")
write.csv(yhat,file="d:/output uji yhat.csv")
}
```

Lampiran 8 :Output untuk Mendapatkan Model Regresi Nonparametrik Spline *Truncated* dengan 1 Knot

No	x_1	x_2	x_3	x_4	x_5	GCV
1	29.54	2.48	11.68	14.57	36.84	77.37
2	31.01	2.66	13.45	16.35	38.15	87.74
3	32.47	2.84	15.22	18.13	39.47	89.89
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
39	85.20	9.27	78.96	82.20	86.84	65.14
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
47	96.91	10.70	93.12	96.44	97.37	81.68
48	98.38	10.88	94.89	98.22	98.68	81.40

Lampiran 9 :Output untuk Mendapatkan Model Regresi Nonparametrik Spline *Truncated* dengan 2 Knot

No	x_1	x_2	x_3	x_4	x_5	GCV
1	28.08	2.30	9.91	12.79	35.52	77.37
	29.54	2.48	11.68	14.57	36.84	
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
1137	80.80	8.74	73.64	76.86	82.89	47.28
	85.20	9.27	78.96	82.20	86.84	
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
1225	98.38	10.88	94.89	98.22	98.68	81.40
	99.84	11.06	96.66	100.00	100.00	

Lampiran 10 :Output untuk Mendapatkan Model Regresi Nonparametrik Spline *Truncated* dengan 3 Knot

No	x_1	x_2	x_3	x_4	x_5	GCV
	29.54	2.48	11.68	14.57	36.84	
1	31.01	2.66	13.45	16.35	38.15	84.72
	32.47	2.84	15.22	18.13	39.47	
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
	61.76	6.41	50.63	53.73	65.79	
14860	73.48	7.84	64.79	67.96	76.31	16.70
	98.38	10.88	94.89	98.22	98.68	
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
	95.45	10.52	91.35	94.66	96.05	
17296	96.91	10.70	93.12	96.44	97.37	135.66
	98.38	10.88	94.89	98.22	98.68	

Lampiran 11 :Output Uji Parameter Spline *Truncated*

```
=========================================
Estimasi Parameter
=========================================
              [,1]
 [1,]   145.1121205
 [2,]     0.6394338
 [3,]    -0.8301846
 [4,]    -0.3882315
 [5,]   218.5828197
 [6,]    -8.8996103
 [7,]    13.9744776
 [8,]   -27.7174617
 [9,] -2182.5298516
[10,]     0.1128036
[11,]     0.4120666
[12,]    -1.8712267
[13,]     8.3082696
[14,]    -5.8222934
[15,]     9.4382433
[16,]    -1.6310158
[17,]   -45.0133333
[18,]     3.0846545
[19,]   -10.0049922
[20,]     5.9472599
[21,]    70.2420654
------------------------------------
Kesimpulan hasil uji serentak
------------------------------------
Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

------------------------------------
Kesimpulan hasil uji individu
------------------------------------
Tolak Ho yakni prediktor signifikan dengan pvalue 5.381968e-07
Tolak Ho yakni prediktor signifikan dengan pvalue 0.001156519
Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.09840962
Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.4265186
Tolak Ho yakni prediktor signifikan dengan pvalue 2.28181e-06
Tolak Ho yakni prediktor signifikan dengan pvalue 0.0003907859
Tolak Ho yakni prediktor signifikan dengan pvalue 0.001072554
Tolak Ho yakni prediktor signifikan dengan pvalue 0.0005397676
Tolak Ho yakni prediktor signifikan dengan pvalue 1.108523e-06
Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.5060771
Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.2628735
```

Lanjutan: Lampiran 11

```
Tolak Ho yakni prediktor signifikan dengan pvalue 0.01794553
Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.1166348
Tolak Ho yakni prediktor signifikan dengan pvalue 1.585132e-06
Tolak Ho yakni prediktor signifikan dengan pvalue 2.270548e-06
Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.08409835
Tolak Ho yakni prediktor signifikan dengan pvalue 0.03876423
Tolak Ho yakni prediktor signifikan dengan pvalue 6.103194e-06
Tolak Ho yakni prediktor signifikan dengan pvalue 1.967112e-05
Tolak Ho yakni prediktor signifikan dengan pvalue 0.0004350001
Tolak Ho yakni prediktor signifikan dengan pvalue 0.03100836
=========================================
nilai t hitung
=========================================
            [,1]
 [1,]  14.3664777
 [2,]   4.9251252
 [3,]  -1.8699826
 [4,]  -0.8376947
 [5,]  11.9021703
 [6,]  -5.8317519
 [7,]   4.9851512
 [8,]  -5.5519032
 [9,] -13.0805264
[10,]   0.6961093
[11,]   1.2043505
[12,]  -2.9672800
[13,]   1.7589405
[14,] -12.4841021
[15,]  11.9099063
[16,]  -1.9718828
[17,]  -2.4691096
[18,]  10.4504147
[19,]  -8.9271679
[20,]   5.7378859
[21,]   2.6125264
Analysis of Variance
========================================
Sumber      df     SS         MS         Fhit
Regresi     20   2363.076   118.1538   25.64944
Error        8   36.85189   4.606486
Total       28   2399.928
========================================
s= 2.146273 Rsq= 98.46446
pvalue(F)= 3.55747e-05
```

Lampiran 12: Tabel Estimasi Parameter Koefisien Regresi Nonparametrik Deret Fourier

Parameter	Estimator	Parameter	Estimator	Parameter	Estimator
β_0	16,88	α_{22}	-4,95	α_{14}	-13,58
b_1	-0,27	α_{32}	-0,04	α_{24}	-9,10
α_{11}	-0,02	b_3	0,41	α_{34}	-9,48
α_{21}	0,80	α_{13}	6,35	b_5	0,14
α_{31}	-3,66	α_{23}	-2,99	α_{15}	4,79
b_2	3,69	α_{33}	0,49	α_{25}	10,28
α_{12}	-1,36	b_4	-0,32	α_{35}	4,55

Lampiran 13: Program Penentuan Nilai K Optimal

```
data=read.table("d://TESIS.txt",header=FALSE)
data
y=as.matrix(data[,1])
x1=as.matrix(data[,2])
x2=as.matrix(data[,3])
x3=as.matrix(data[,4])
x4=as.matrix(data[,5])
x5=as.matrix(data[,6])
MPL<-function(x,eps=1e-0009)
{
    x<-as.matrix(x)
    xsvd<-svd(x)
    diago<-xsvd$d[xsvd$d>eps]
    if(length(diago)==1)
    {
      xplus<-as.matrix(xsvd$v[,1])%*%t(as.matrix(xsvd$u[,1])/diago)
    }
      else
      {
          xplus<-
          xsvd$v[,1:length(diago)]%*%diag(1/diago)%*%t(xsvd$u[,1:len
          gth(diago)])
      }
      return(xplus)
}
deretfourier<-function(y, x1, x2, x3, x4, x5, K)
{
    N=length(y)
    a<-(5*(K+1))+1
    C<-matrix(0,N,a)
    hasil<-matrix(0,K,2)
    for (k in 1:K)
    {
      for(i in 1:N)
      {
            for(j in 1:k)
            {
                C[i,1]=1
                C[i,2]<-x1[i]
                C[i,2+j]=cos(j*x1[i])
                C[i,3+k]<-x2[i]
                C[i,3+k+j]<-cos(j*x2[i])
                C[i,4+(2*k)]<-x3[i]
                C[i,4+(2*k)+j]<-cos(j*x3[i])
                C[i,5+(3*k)]<-x4[i]
                C[i,5+(3*k)+j]<-cos(j*x4[i])
```

Lanjutan: Lampiran 13

```
                C[i,6+(4*k)]<-x5[i]
                C[i,6+(4*k)+j]<-cos(j*x5[i])
              }
          }
          I<-diag(1,N,N)
          A<-C%*%MPL(t(C)%*%C)%*%t(C)
          Atot=A
          ytop<-Atot%*%y
          df<-sum(diag(Atot))
          W<-(1/N)*I
          atas<-t(y-ytop)%*%W%*%(y-ytop)
          bawah<-((1-df)/N)^2
          GCV<-atas/bawah
          hasil[k,1]<-k
          hasil[k,2]<-GCV
         }
print(hasil)
GCV2<-min(hasil[,2])
s<-1
repeat{
if(hasil[s,2]==GCV2)
{
  kOpt<-hasil[s,1]
  GCVOpt<-GCV2
  break
}
  else s<-s+1
}
cat("nilai K optimal adalah \t",kOpt,"\n")
print(C)
}
```

Lampiran 14: Program untuk Estimasi Model

```
data=read.table("d://TESIS.txt",header=FALSE)
data
y=as.matrix(data[,1])
x1=as.matrix(data[,2])
x2=as.matrix(data[,3])
x3=as.matrix(data[,4])
x4=as.matrix(data[,5])
x5=as.matrix(data[,6])
MPL<-function(x,eps=1e-009)
{
    x<-as.matrix(x)
    xsvd<-svd(x)
    diago<-xsvd$d[xsvd$d>eps]
    if(length(diago)==1)
    {
      xplus<-as.matrix(xsvd$v[,1])%*%t(as.matrix(xsvd$u[,1])/diago)
    }
      else
      {
         xplus<-
         xsvd$v[,1:length(diago)]%*%diag(1/diago)%*%t(xsvd$u[,1:
         length(diago)])
      }
      return(xplus)
    }
    estimasi<-function(y, x1, x2, x3, x4, x5,kOptimal)
    {
      N<-length(y)
      X<-matrix(0,N,2)
      k<-kOptimal
      a<-(5*(k+1))+1
      C<-matrix(0,N,a)
      hasil<-matrix(0,k,2)
      error<-rep(0,N)
cat("=============================================================”)
cat("\nx\ty\t\tytopi\t\terror")
cat("\n===========================================================\n")
    for(i in 1:N)
    {
      for(j in 1:k)
      {
        C[i,1]=1
        C[i,2]<-x1[i]
        C[i,2+j]=cos(j*x1[i])
        C[i,3+k]<-x2[i]
        C[i,3+k+j]<-cos(j*x2[i])
        C[i,4+(2*k)]<-x3[i]
        C[i,4+(2*k)+j]<-cos(j*x3[i])
        C[i,5+(3*k)]<-x4[i]
        C[i,5+(3*k)+j]<-cos(j*x4[i])
```

Lanjutan: Lampiran 14.

```
        C[i,6+(4*k)]<-x5[i]
        C[i,6+(4*k)+j]<-cos(j*x5[i])
    }
}
I<-diag(1,N,N)
A<-C%*%MPL(t(C)%*%C)%*%t(C)
Atot<-A
ytop<-Atot%*%y
alfa<-rep(0,a)
alfa<-MPL(t(C)%*%C)%*%t(C)%*%y
error<-y-ytop
MSE<-sum((error)^2)/N
for(i in 1:N)
{
    cat("\n","\t",y[i],"\t",ytop[i],"\t",error[i],"\n")
}
cat("\n===========================================================\n")
cat("\n MSE=",MSE,"\n")
Q<-0
    for(i in 1:N)
    {
       q<-(ytop[i]-mean(y))^2
       Q<-Q+q
    }
    F<-0
    for(i in 1:N)
    {
       f<-(y[i]-mean(y))^2
       F<-F+f
    }
R<-Q/F
cat("Nilai Koefisien Determinasi = ")
print(R)
print(alfa)
}
```

Lampiran 15. Nilai GCV, R^2 dan MSE pada model Deret Fourier

K	GCV	R^2	MSE
1	214,27	69,21%	25,48
2	84,73	72,61%	22,67
3	18,79	89,20%	8,94

(...Halaman Ini Sengaja Dikosongkan...)